Accelerating Sustainable Glass Discovery: Integrating Molecular Dynamics, Machine Learning, and Robotic Synthesis

Felix Arendt[a], Tina Waurischk[b], Stefan Reinsch[b], Andrea S. S. de Camargo,[b,c] Marek Sierka[a]*

[a] *Otto Schott Institute of Materials Research, Friedrich-Schiller University Jena, Löbdergraben 32, 07743 Jena, Germany*

[b] *Federal Institute for Materials Research and Testing (BAM), Richard-Willstätter-Strasse 11, 12489 Berlin, Germany*

[c] *Otto Schott Institute of Materials Research, Friedrich-Schiller University Jena, Lessingstrasse 12-14, 07743 Jena, Germany*

E-mail: marek.sierka@uni-jena.de

Supplementary information

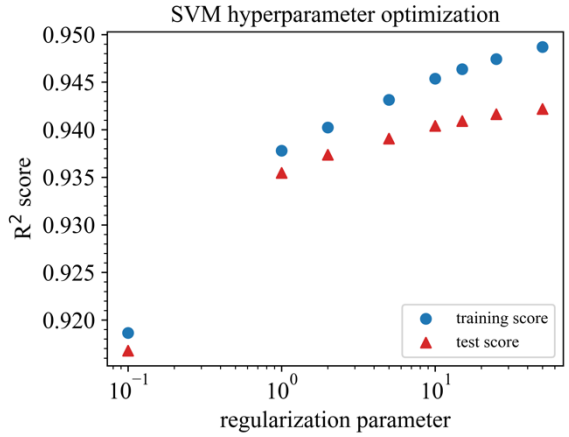

Figure S1: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.01$, descriptor set AC, and density as target property.

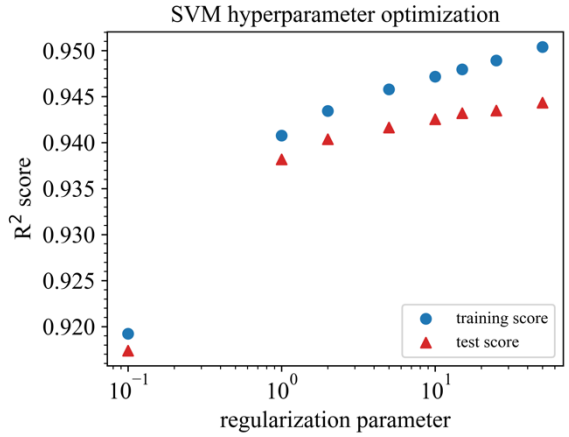

Figure S2: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.1$, descriptor set AC, and density as target property.

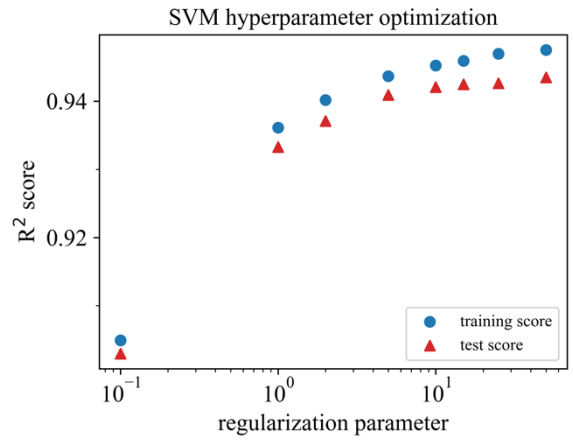

Figure S3: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, ϵ = 0.2, descriptor set AC, and density as target property.

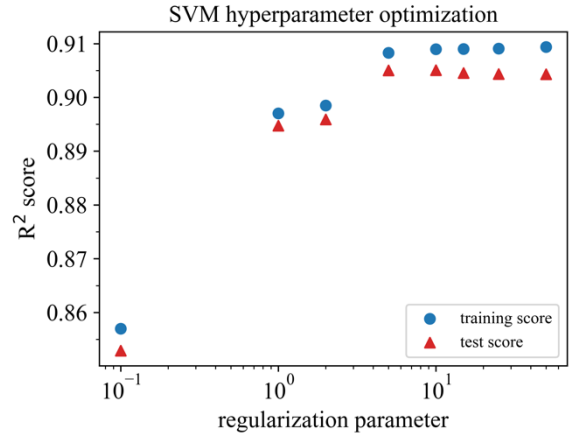

Figure S4: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, ϵ = 0.5, descriptor set AC, and density as target property.

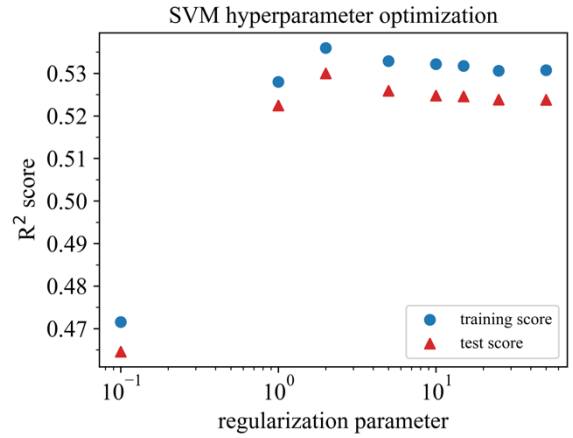

Figure S5: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, ϵ = 1.0, descriptor set AC, and density as target property.

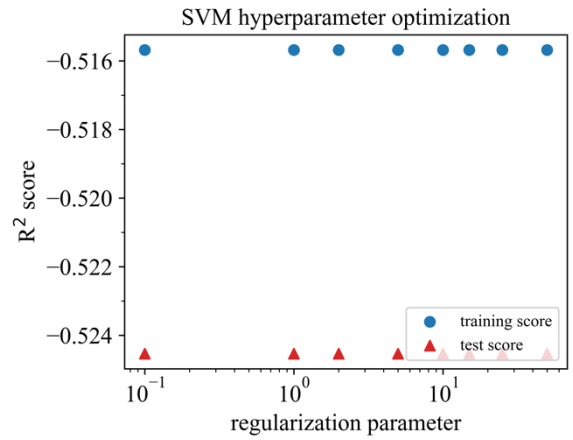

Figure S6: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, ϵ = 2.0, descriptor set AC, and density as target property.

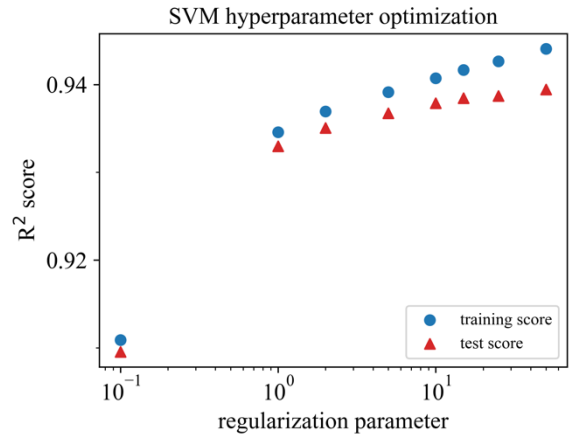

Figure S7: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, ϵ = 0.01, descriptor set A, and density as target property.

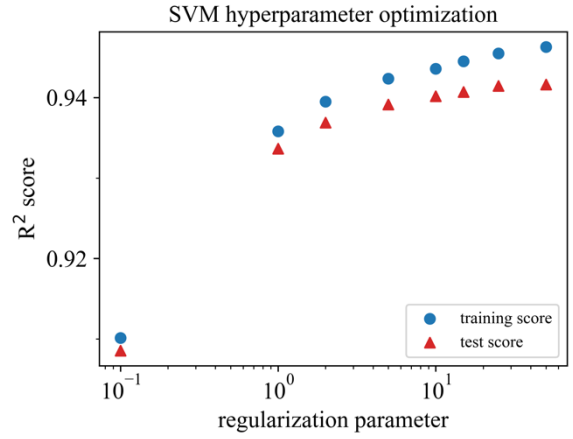

Figure S8: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, ϵ = 0.1, descriptor set A, and density as target property.

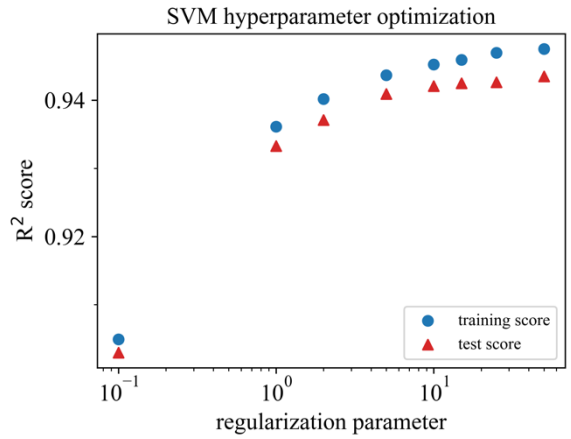

Figure S9: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.2$, descriptor set A, and density as target property.

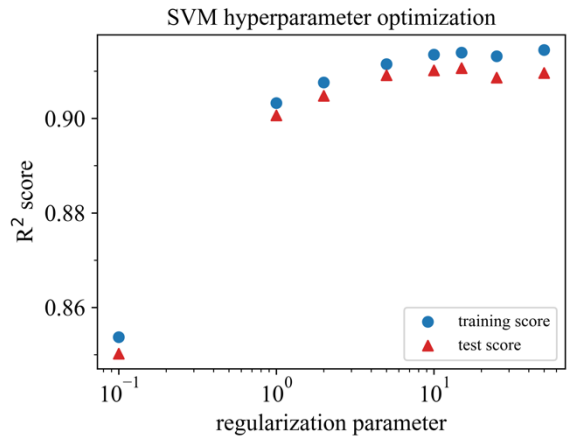

Figure S10: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.5$, descriptor set A, and density as target property.

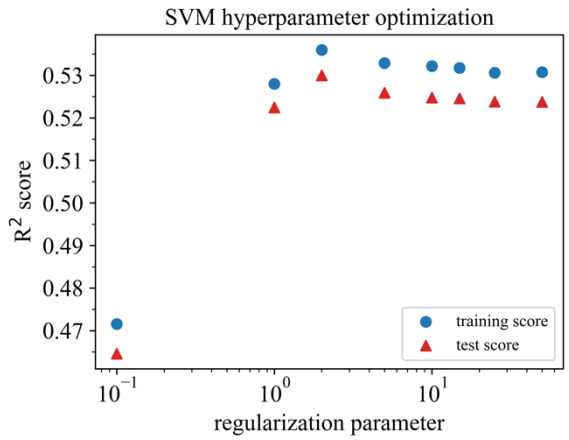

Figure S11: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 1.0$, descriptor set A, and density as target property.

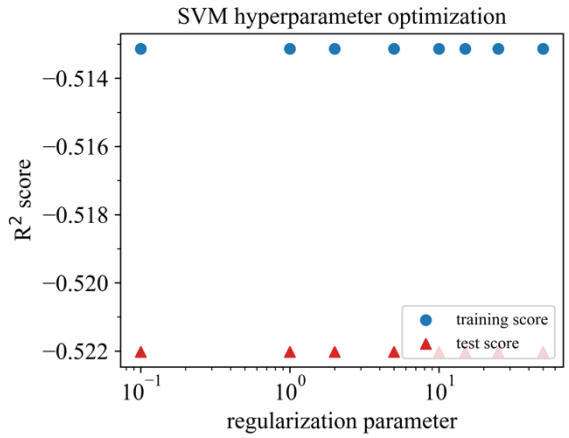

Figure S12: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 2.0$, descriptor set A, and density as target property.

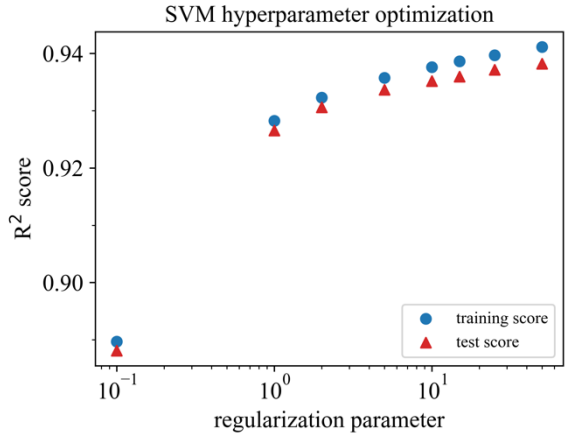

Figure S13: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.01$, descriptor set C, and density as target property.

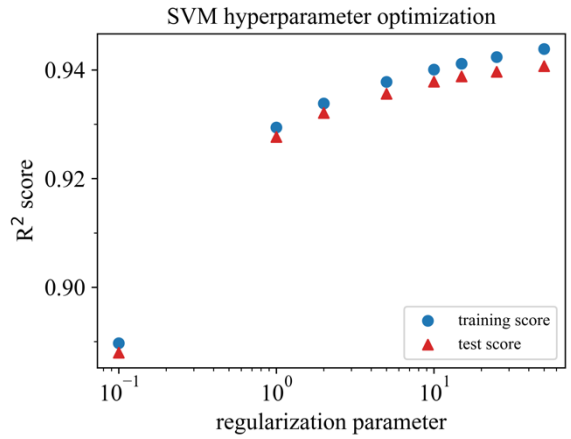

Figure S14: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.1$, descriptor set C, and density as target property.

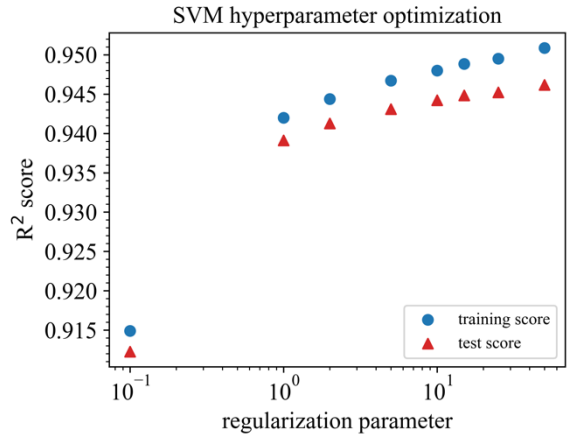

Figure S15: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.2$, descriptor set C, and density as target property.

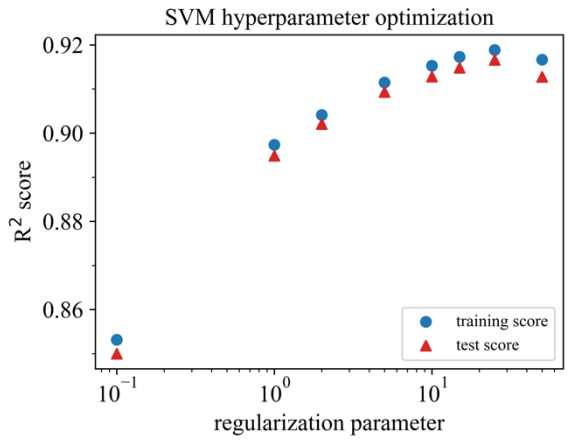

Figure S16: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.5$, descriptor set C, and density as target property.

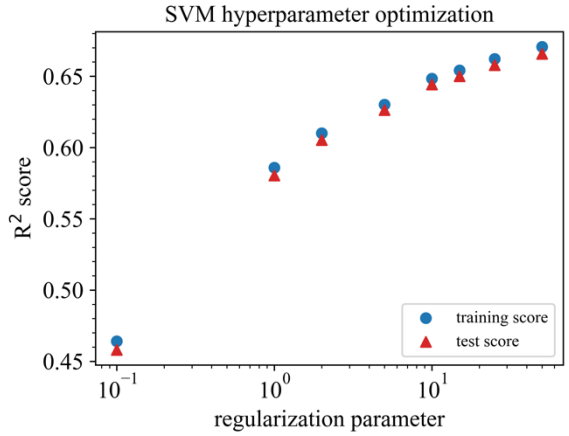

Figure S17: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 1.0$, descriptor set C, and density as target property.

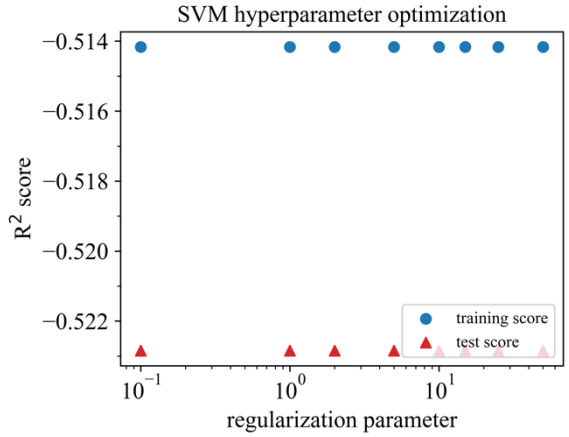

Figure S18: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 2.0$, descriptor set C, and density as target property.

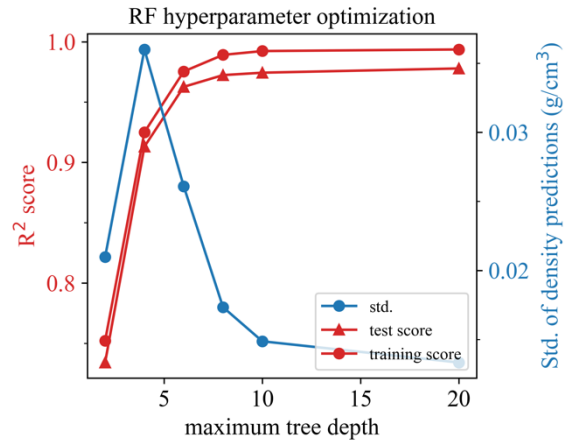

Figure S19: Hyperparameter optimization for RF trained with the V18 dataset, 4 estimators, descriptor set AC, and density as target property.

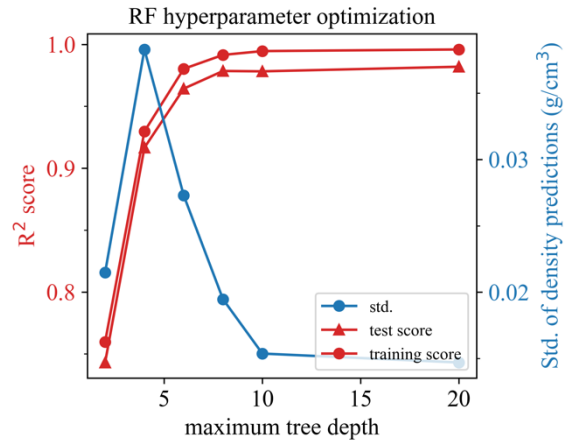

Figure S20: Hyperparameter optimization for RF trained with the V18 dataset, 8 estimators, descriptor set AC, and density as target property.

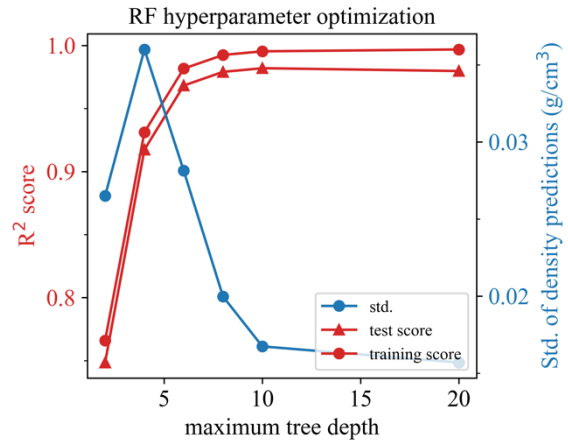

Figure S21: Hyperparameter optimization for RF trained with the V18 dataset, 16 estimators, descriptor set AC, and density as target property.

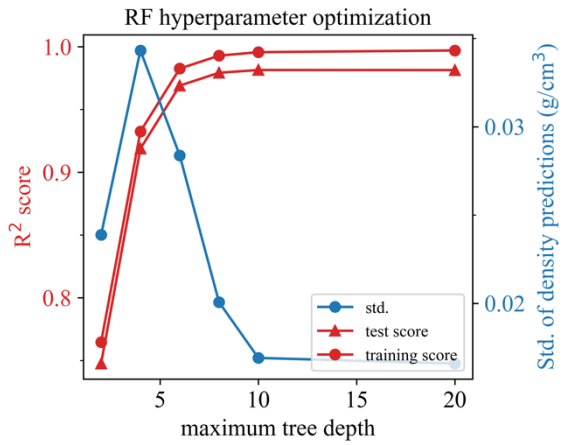

Figure S22: Hyperparameter optimization for RF trained with the V18 dataset, 32 estimators, descriptor set AC, and density as target property.

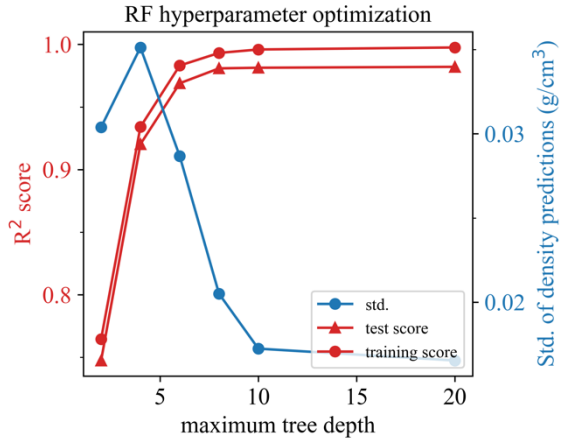

Figure S23: Hyperparameter optimization for RF trained with the V18 dataset, 64 estimators, descriptor set AC, and density as target property.

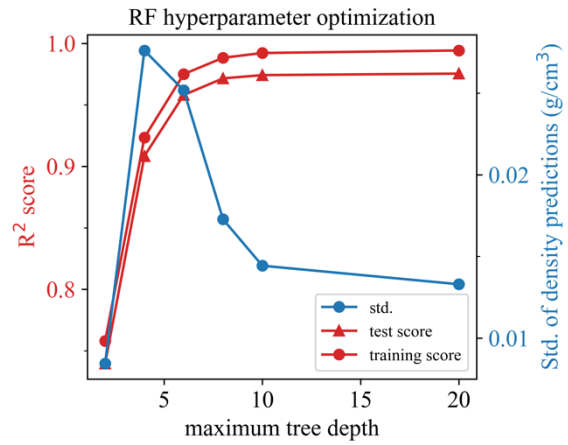

Figure S24: Hyperparameter optimization for RF trained with the V18 dataset, 4 estimators, descriptor set A, and density as target property.

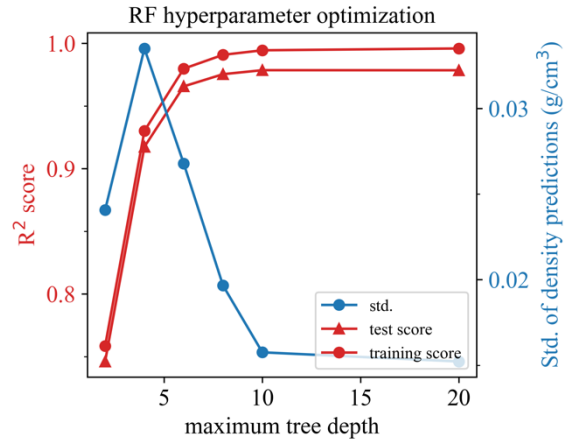

Figure S25: Hyperparameter optimization for RF trained with the V18 dataset, 8 estimators, descriptor set A, and density as target property.

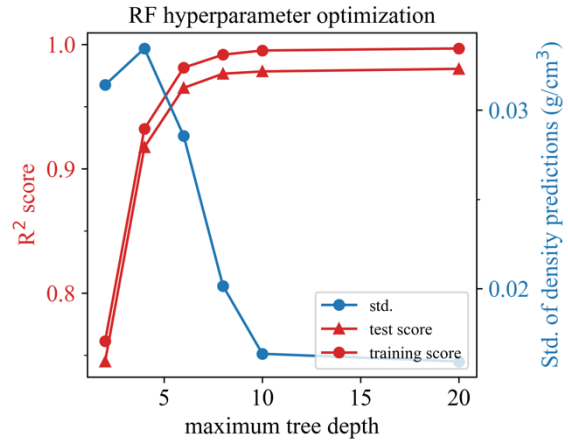

Figure S26: Hyperparameter optimization for RF trained with the V18 dataset, 16 estimators, descriptor set A, and density as target property.

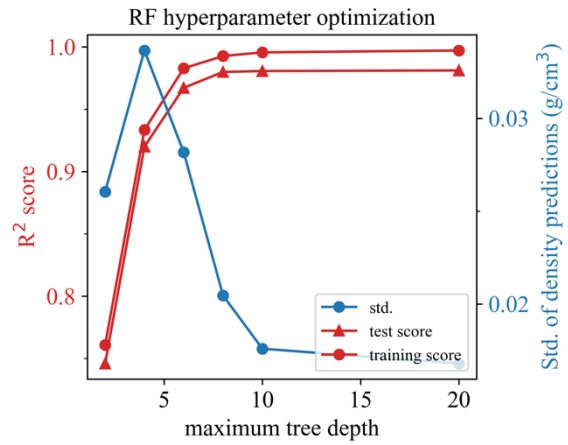

Figure S27: Hyperparameter optimization for RF trained with the V18 dataset, 32 estimators, descriptor set A, and density as target property.

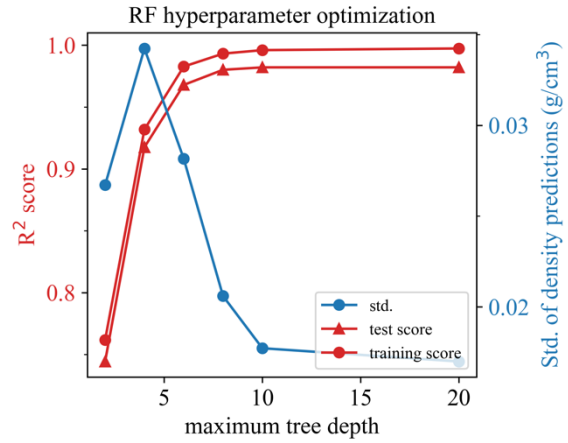

Figure S28: Hyperparameter optimization for RF trained with the V18 dataset, 64 estimators, descriptor set A, and density as target property.

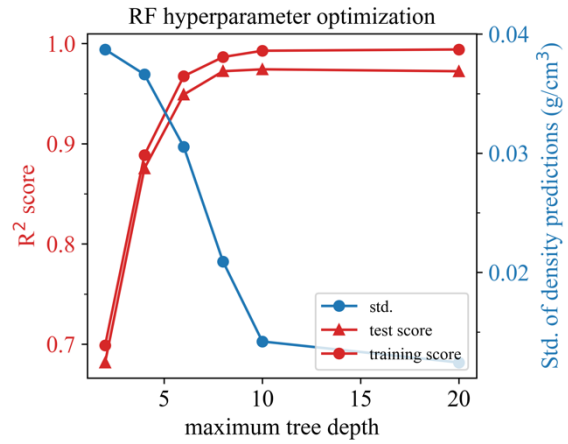

Figure S29: Hyperparameter optimization for RF trained with the V18 dataset, 4 estimators, descriptor set C, and density as target property.

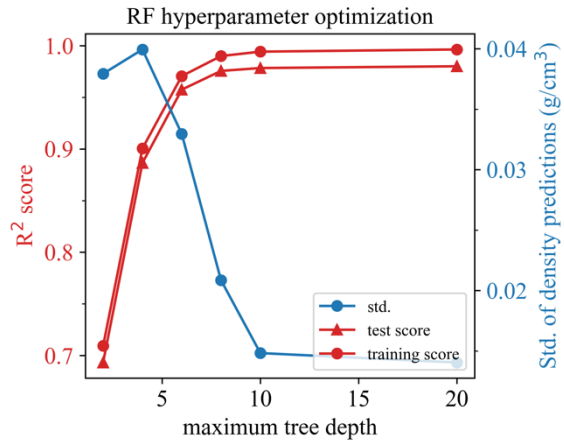

Figure S30: Hyperparameter optimization for RF trained with the V18 dataset, 8 estimators, descriptor set C, and density as target property.

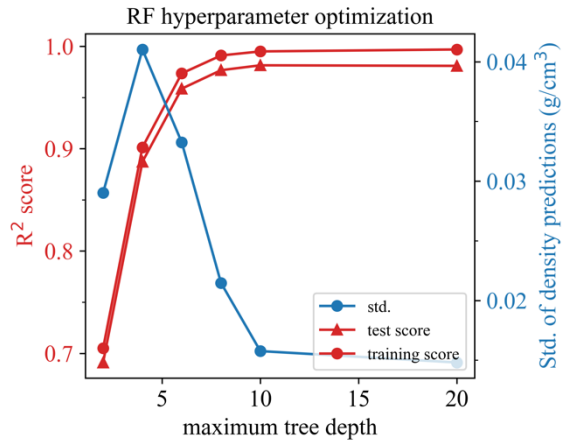

Figure S31: Hyperparameter optimization for RF trained with the V18 dataset, 16 estimators, descriptor set C, and density as target property.

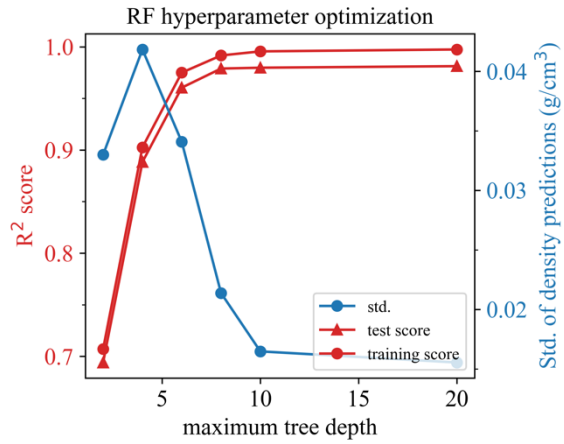

Figure S32: Hyperparameter optimization for RF trained with the V18 dataset, 32 estimators, descriptor set C, and density as target property.

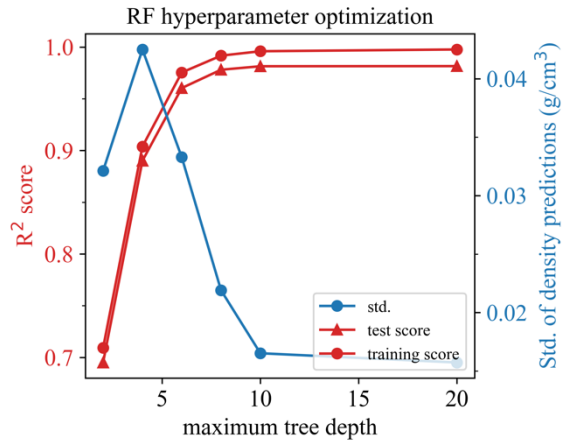

Figure S33: Hyperparameter optimization for RF trained with the V18 dataset, 64 estimators, descriptor set C, and density as target property.

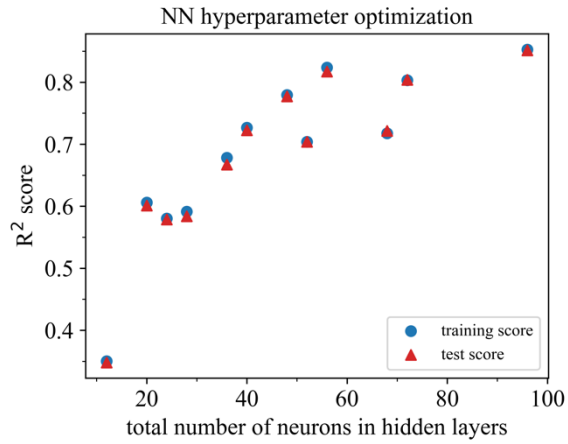

Figure S34: Hyperparameter optimization for NN trained with the V18 dataset, descriptor set AC, and density as target property.

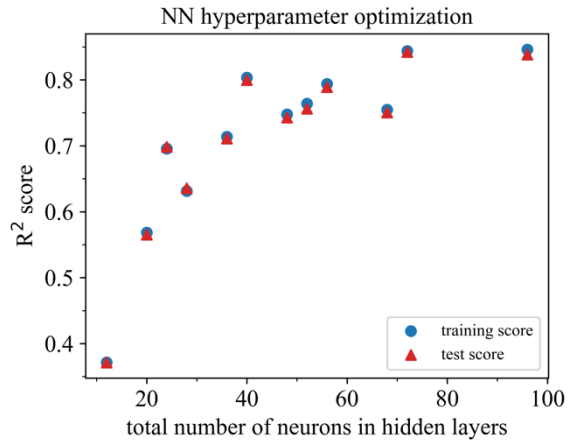

Figure S35: Hyperparameter optimization for NN trained with the V18 dataset, descriptor set A, and density as target property.

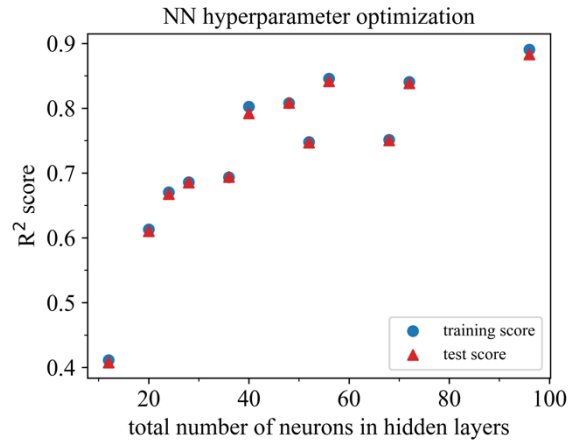

Figure S36: Hyperparameter optimization for NN trained with the V18 dataset, descriptor set C, and density as target property.

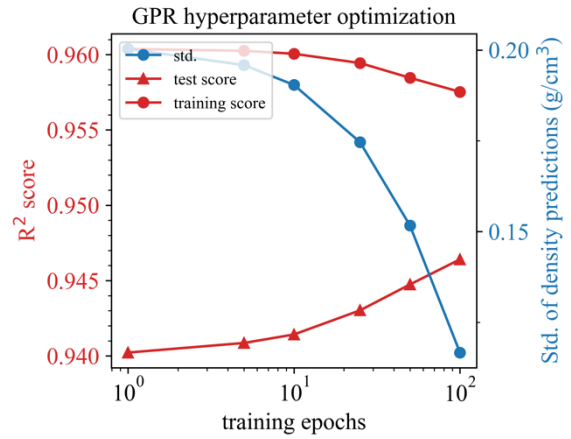

Figure S37: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.01, descriptor set AC, and density as target property.

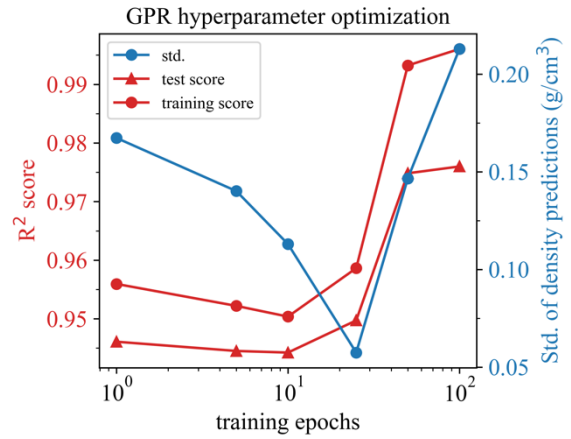

Figure S38: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.1, descriptor set AC, and density as target property.

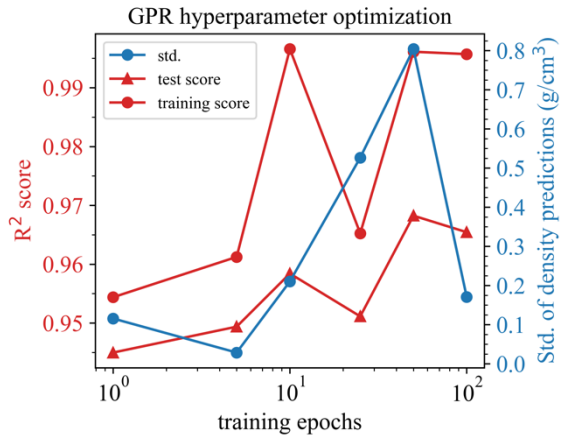

Figure S39: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 1.0, descriptor set AC, and density as target property.

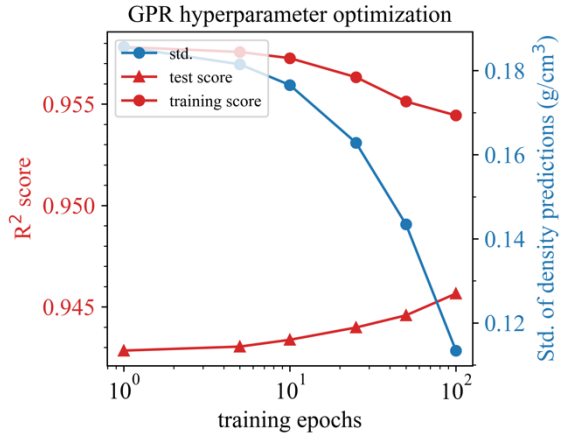

Figure S40: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.01, descriptor set A, and density as target property.

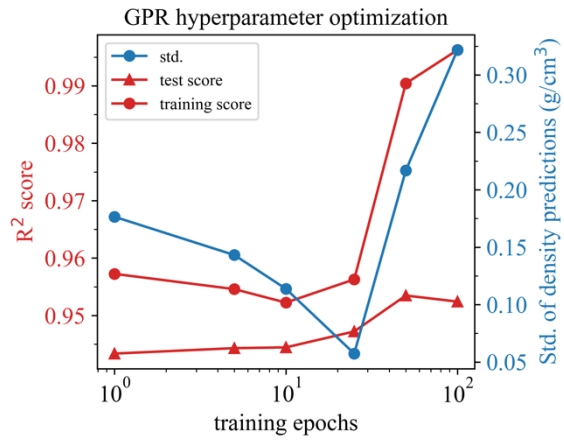

Figure S41: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.1, descriptor set A, and density as target property.

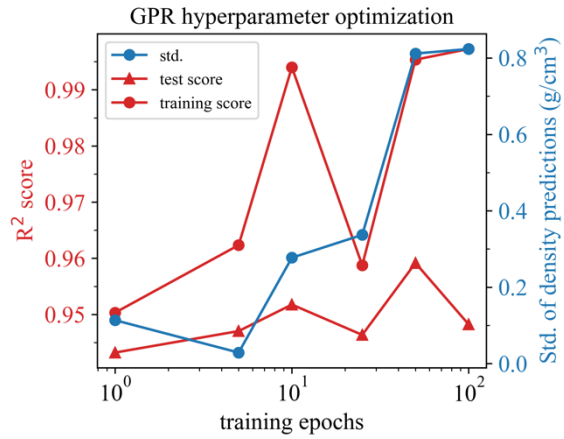

Figure S42: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 1.0, descriptor set A, and density as target property.

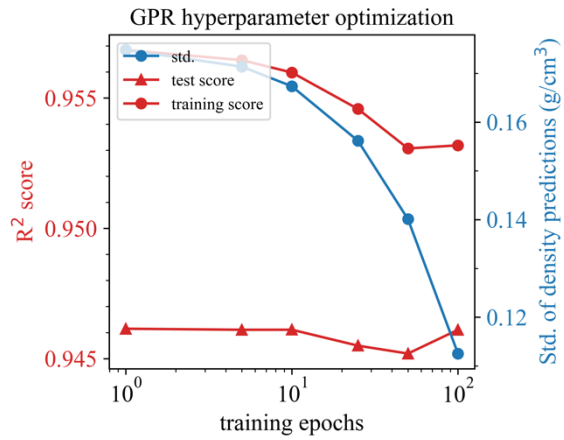

Figure S43: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.01, descriptor set C, and density as target property.

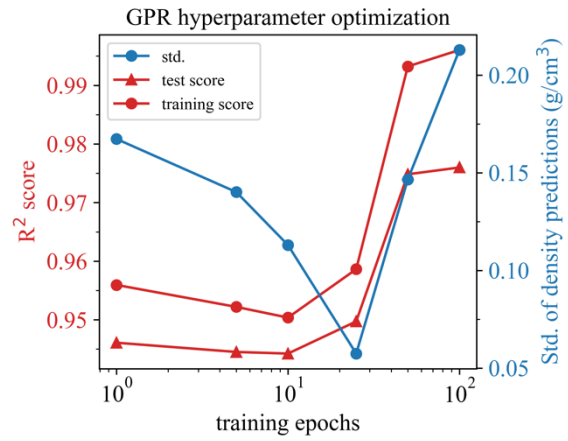

Figure S44: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.1, descriptor set C, and density as target property.

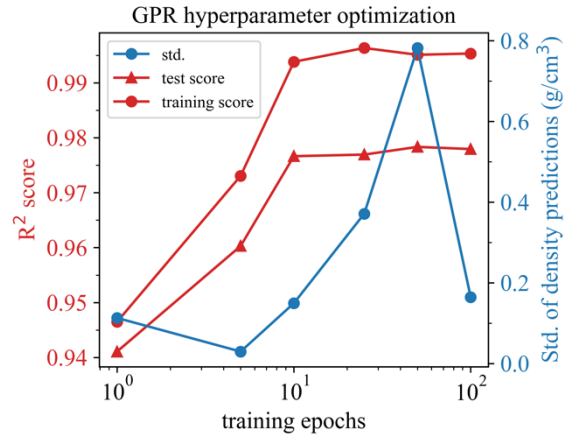

Figure S45: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 1.0, descriptor set C, and density as target property.

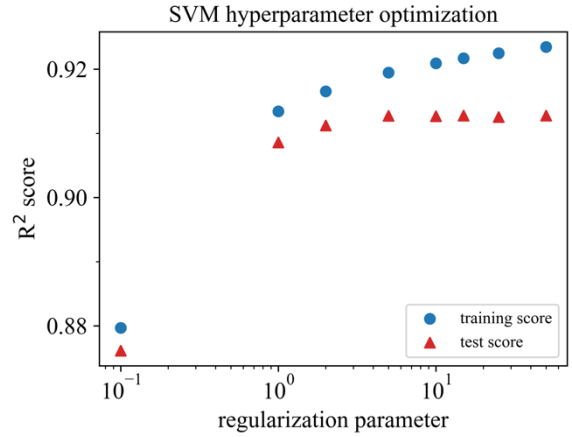

Figure S46: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, ϵ = 0.01, descriptor set AC, and bulk modulus as target property.

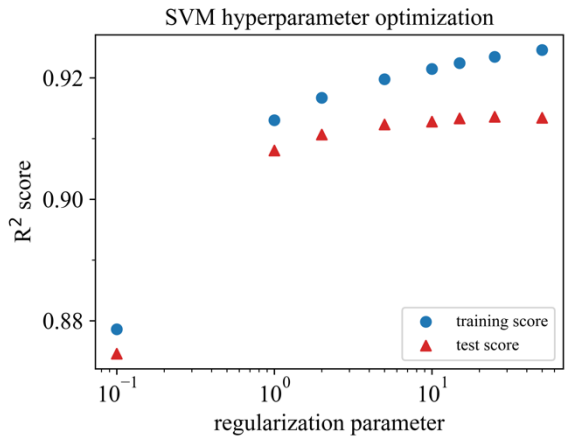

Figure S47: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.1$, descriptor set AC, and bulk modulus as target property.

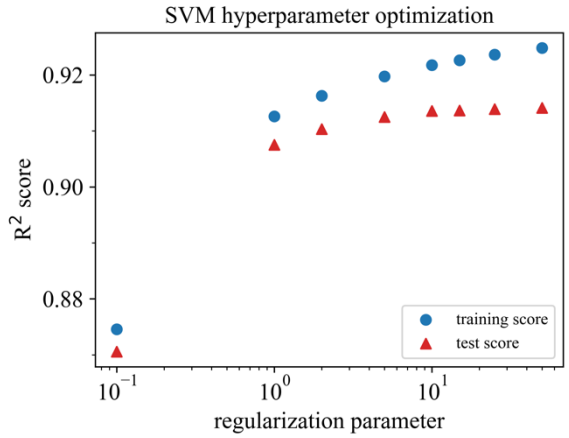

Figure S48: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.2$, descriptor set AC, and bulk modulus as target property.

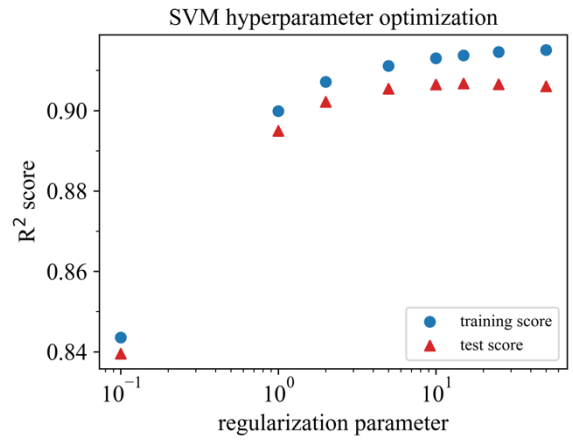

Figure S49: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.5$, descriptor set AC, and bulk modulus as target property.

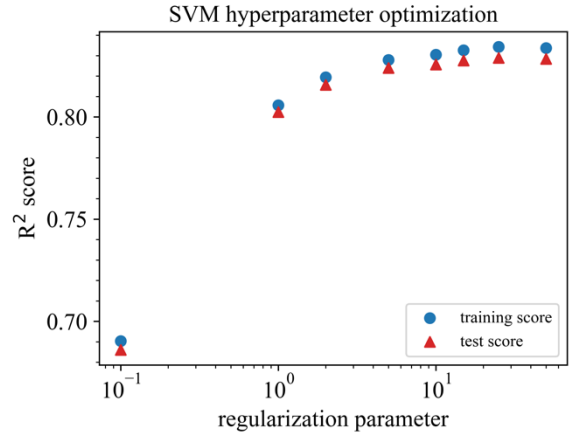

Figure S50: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 1.0$, descriptor set AC, and bulk modulus as target property.

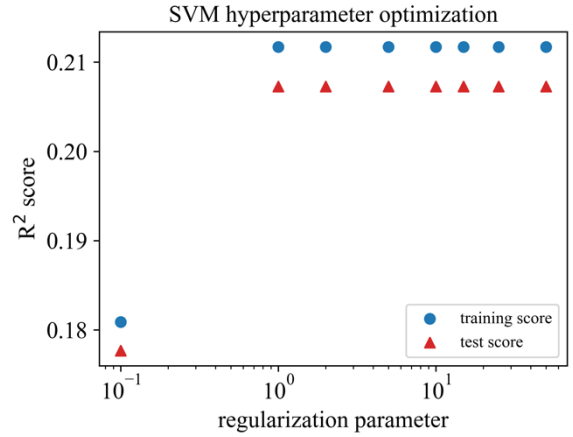

Figure S51: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 2.0$, descriptor set AC, and bulk modulus as target property.

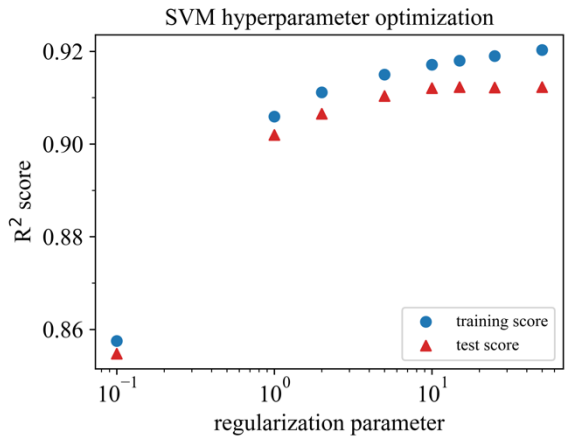

Figure S52: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.01$, descriptor set A, and bulk modulus as target property.

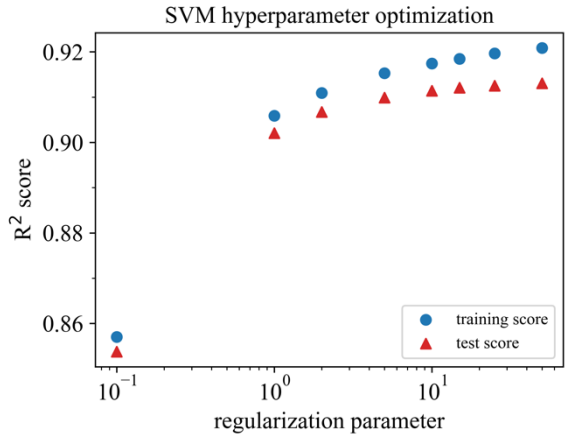

Figure S53: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.1$, descriptor set A, and bulk modulus as target property.

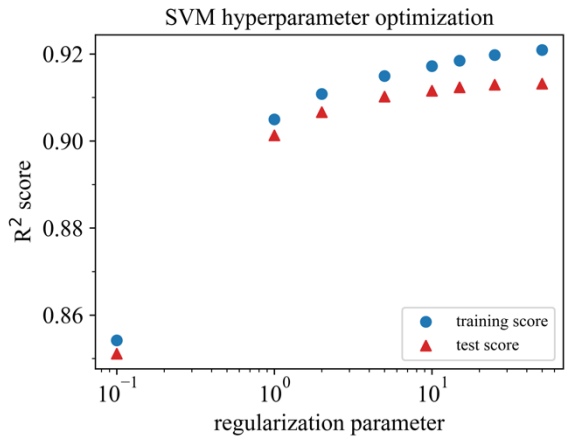

Figure S54: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.2$, descriptor set A, and bulk modulus as target property.

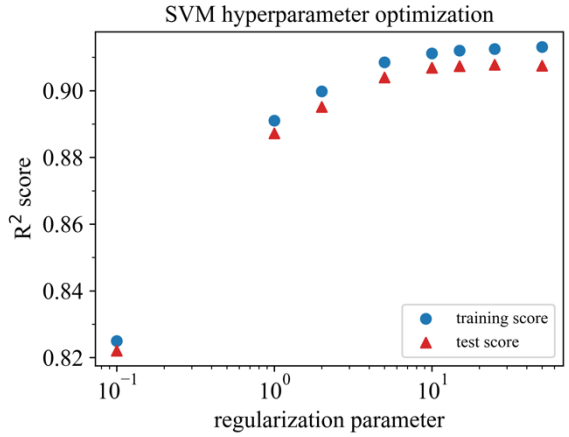

Figure S55: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.5$, descriptor set A, and bulk modulus as target property.

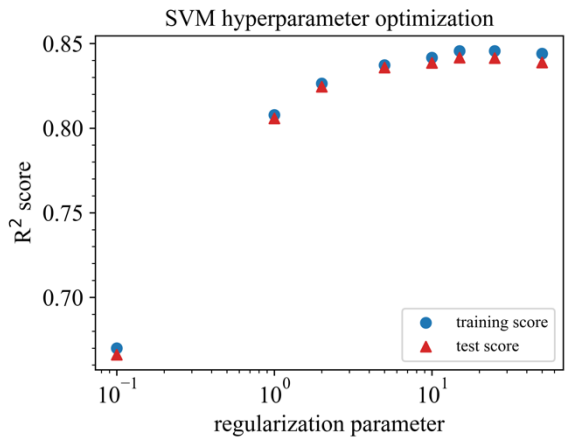

Figure S56: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 1.0$, descriptor set A, and bulk modulus as target property.

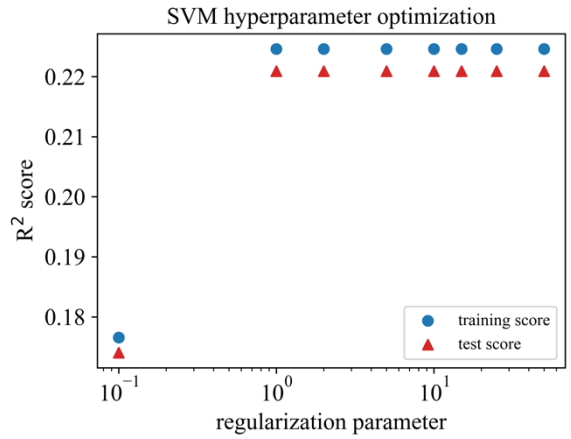

Figure S57: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 2.0$, descriptor set A, and bulk modulus as target property.

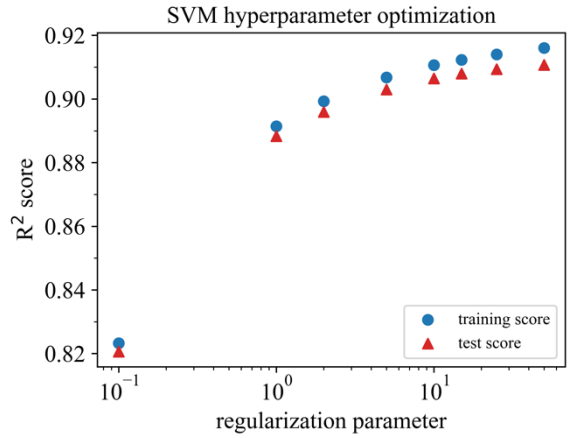

Figure S58: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.01$, descriptor set C, and bulk modulus as target property.

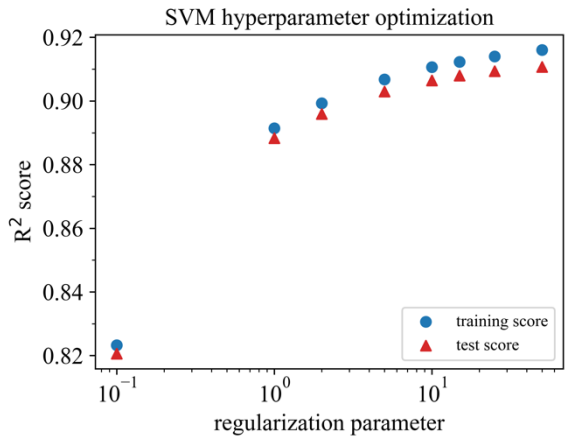

Figure S59: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.1$, descriptor set C, and bulk modulus as target property.

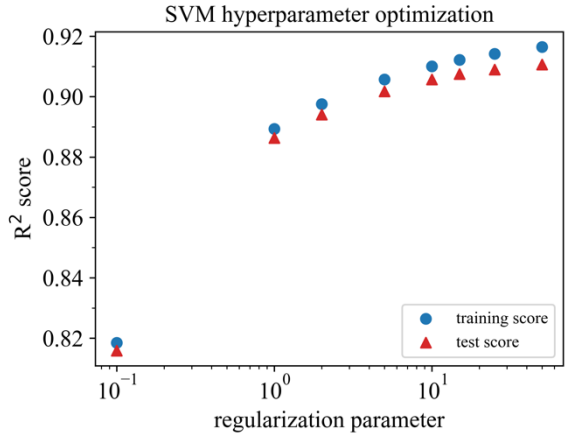

Figure S60: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.2$, descriptor set C, and bulk modulus as target property.

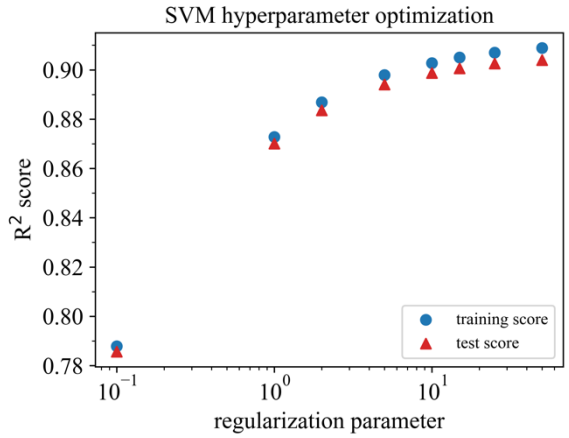

Figure S61: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.5$, descriptor set C, and bulk modulus as target property.

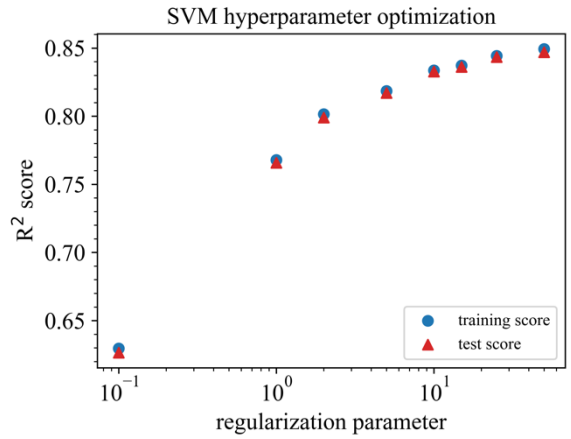

Figure S62: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 1.0$, descriptor set C, and bulk modulus as target property.

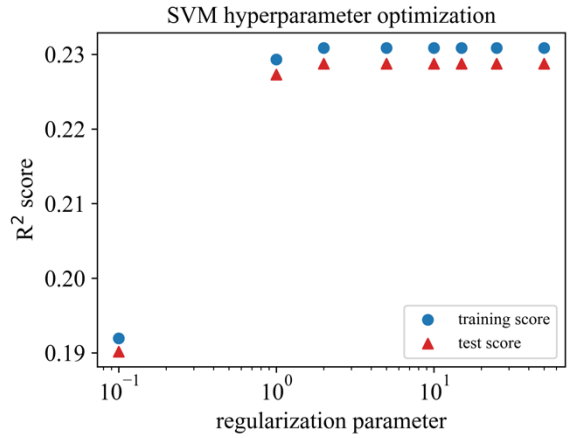

Figure S63: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 2.0$, descriptor set C, and bulk modulus as target property.

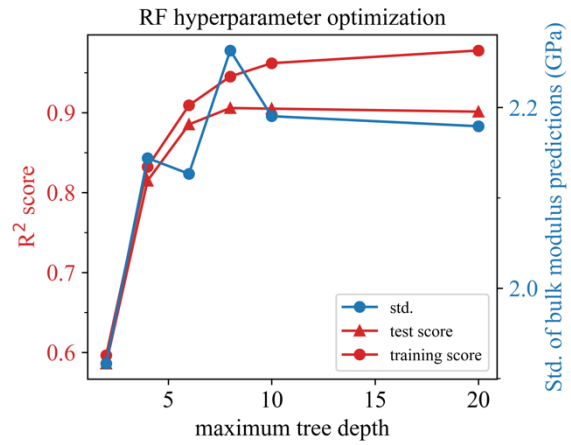

Figure S64: Hyperparameter optimization for RF trained with the V18 dataset, 4 estimators, descriptor set AC, and bulk modulus as target property.

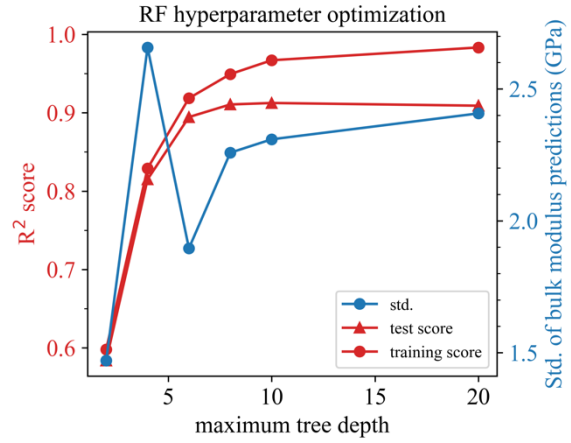

Figure S65: Hyperparameter optimization for RF trained with the V18 dataset, 8 estimators, descriptor set AC, and bulk modulus as target property.

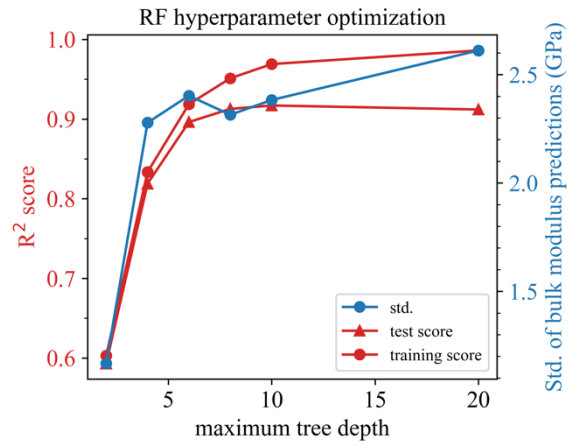

Figure S66: Hyperparameter optimization for RF trained with the V18 dataset, 16 estimators, descriptor set AC, and bulk modulus as target property.

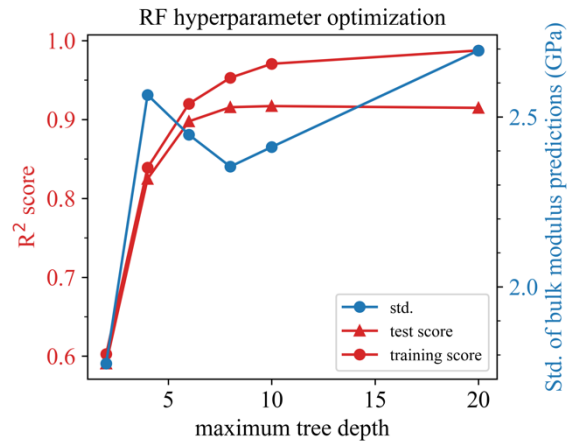

Figure S67: Hyperparameter optimization for RF trained with the V18 dataset, 32 estimators, descriptor set AC, and bulk modulus as target property.

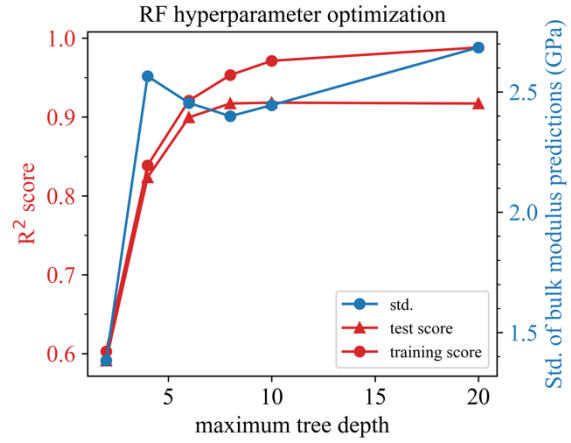

Figure S68: Hyperparameter optimization for RF trained with the V18 dataset, 64 estimators, descriptor set AC, and bulk modulus as target property.

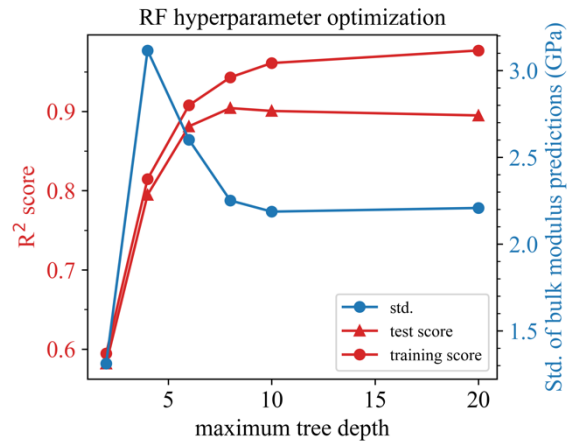

Figure S69: Hyperparameter optimization for RF trained with the V18 dataset, 4 estimators, descriptor set A, and bulk modulus as target property.

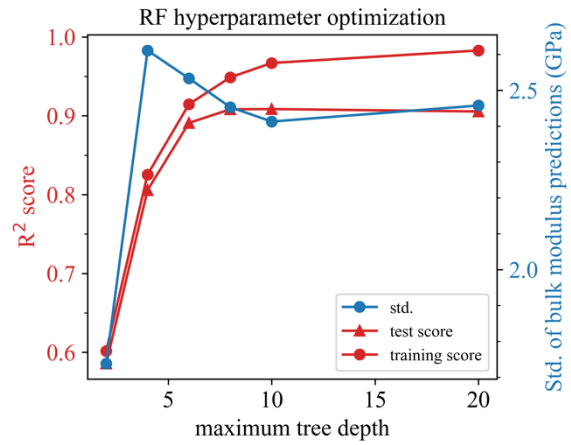

Figure S70: Hyperparameter optimization for RF trained with the V18 dataset, 8 estimators, descriptor set A, and bulk modulus as target property.

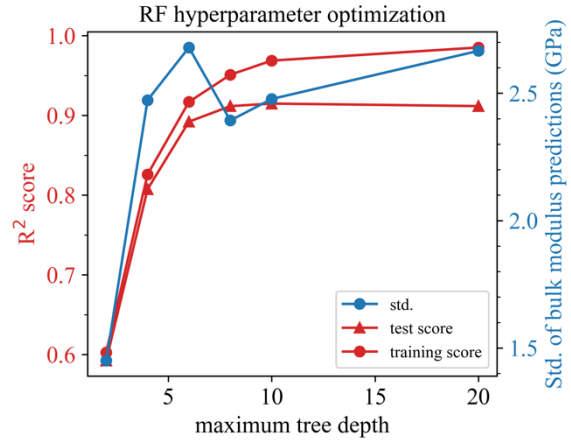

Figure S71: Hyperparameter optimization for RF trained with the V18 dataset, 16 estimators, descriptor set A, and bulk modulus as target property.

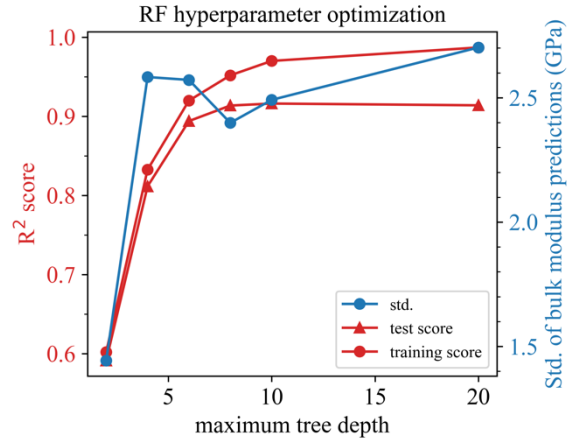

Figure S72: Hyperparameter optimization for RF trained with the V18 dataset, 32 estimators, descriptor set A, and bulk modulus as target property.

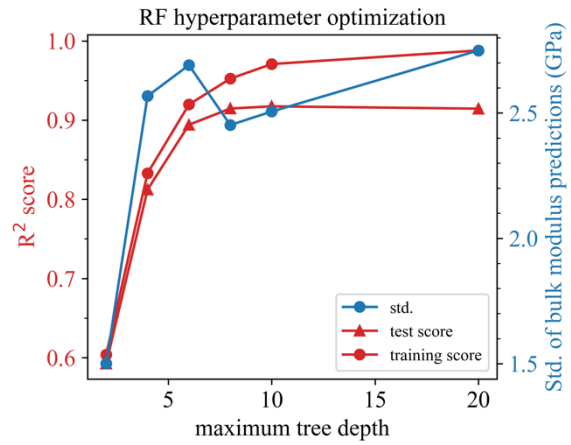

Figure S73: Hyperparameter optimization for RF trained with the V18 dataset, 64 estimators, descriptor set A, and bulk modulus as target property.

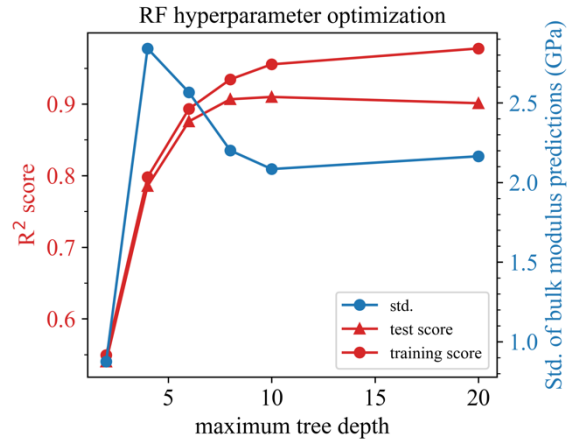

Figure S74: Hyperparameter optimization for RF trained with the V18 dataset, 4 estimators, descriptor set C, and bulk modulus as target property.

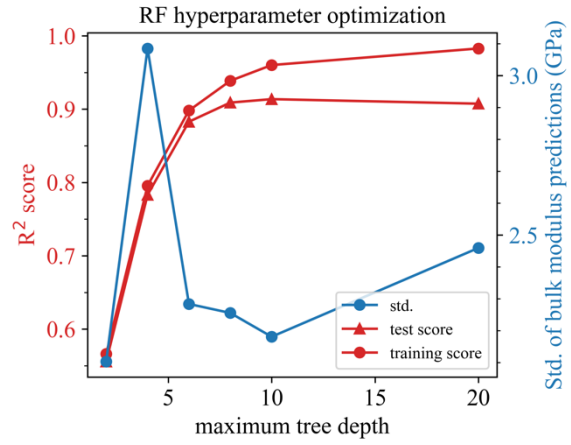

Figure S75: Hyperparameter optimization for RF trained with the V18 dataset, 8 estimators, descriptor set C, and bulk modulus as target property.

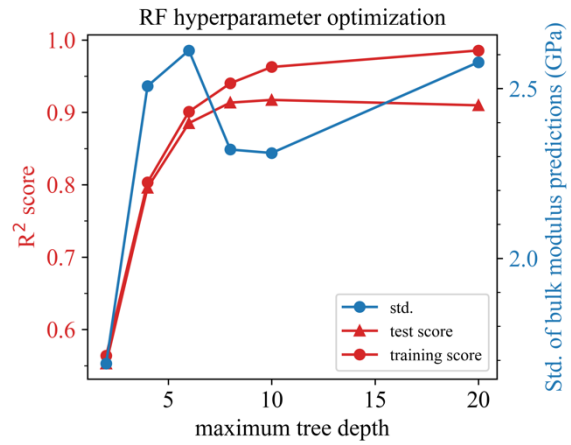

Figure S76: Hyperparameter optimization for RF trained with the V18 dataset, 16 estimators, descriptor set C, and bulk modulus as target property.

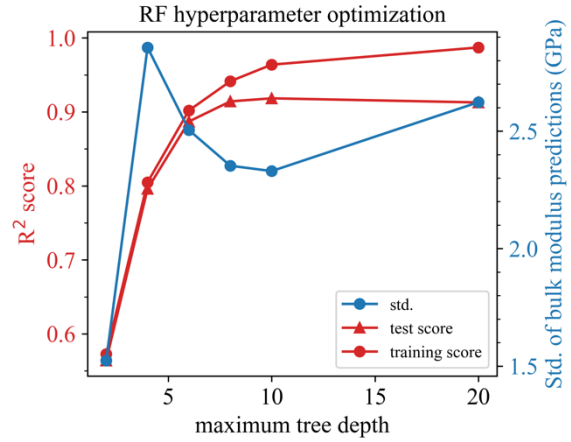

Figure S77: Hyperparameter optimization for RF trained with the V18 dataset, 32 estimators, descriptor set C, and bulk modulus as target property.

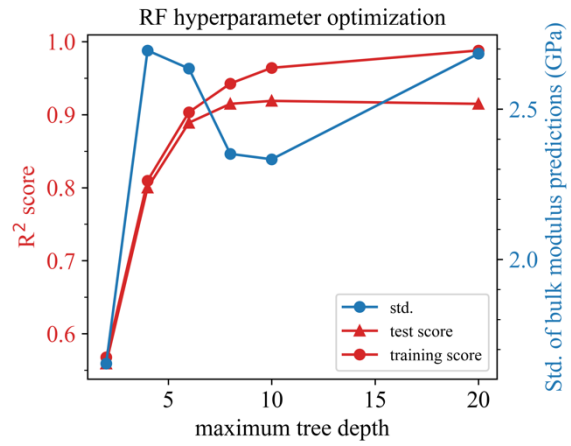

Figure S78: Hyperparameter optimization for RF trained with the V18 dataset, 64 estimators, descriptor set C, and bulk modulus as target property.

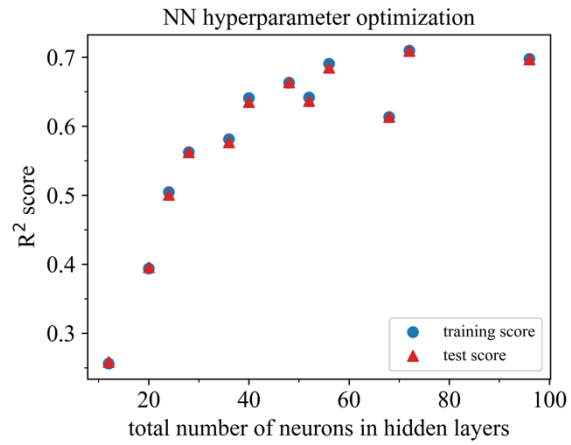

Figure S79: Hyperparameter optimization for NN trained with the V18 dataset, descriptor set AC, and bulk modulus as target property.

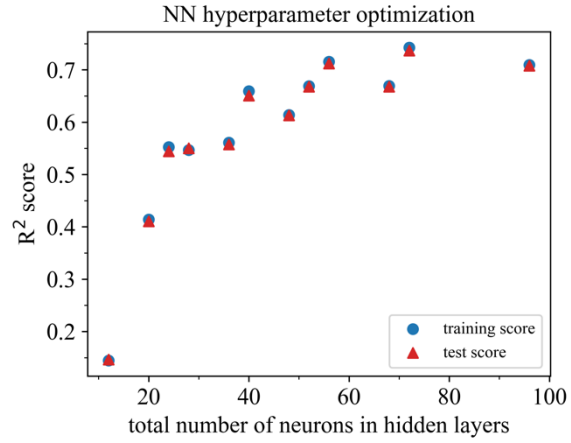

Figure S80: Hyperparameter optimization for NN trained with the V18 dataset, descriptor set A, and bulk modulus as target property.

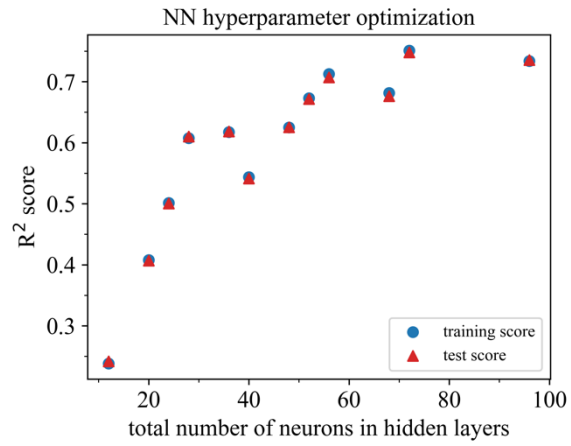

Figure S81: Hyperparameter optimization for NN trained with the V18 dataset, descriptor set C, and bulk modulus as target property.

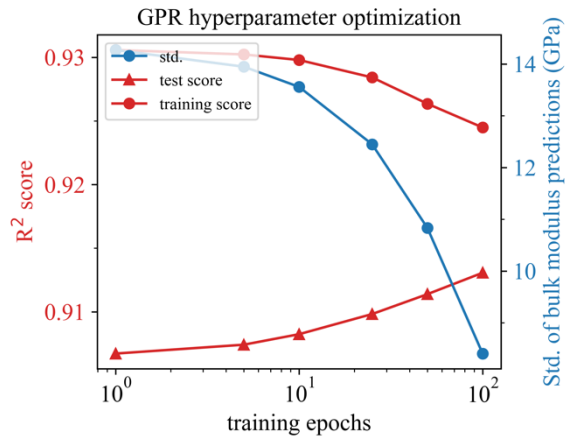

Figure S82: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.01, descriptor set AC, and bulk modulus as target property.

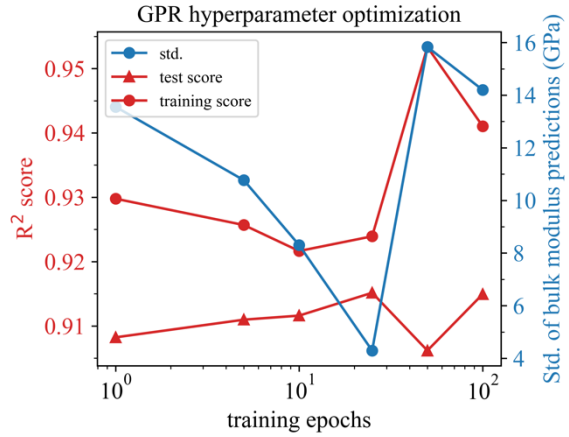

Figure S83: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.1, descriptor set AC, and bulk modulus as target property.

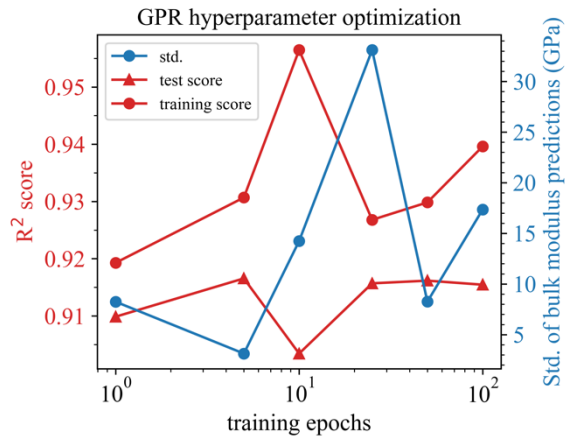

Figure S84: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 1.0, descriptor set AC, and bulk modulus as target property.

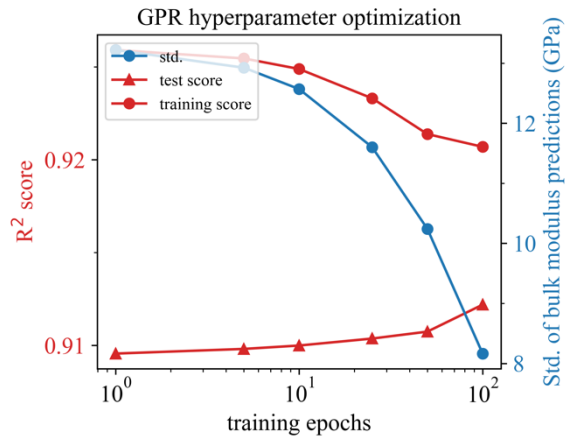

Figure S85: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.01, descriptor set A, and bulk modulus as target property.

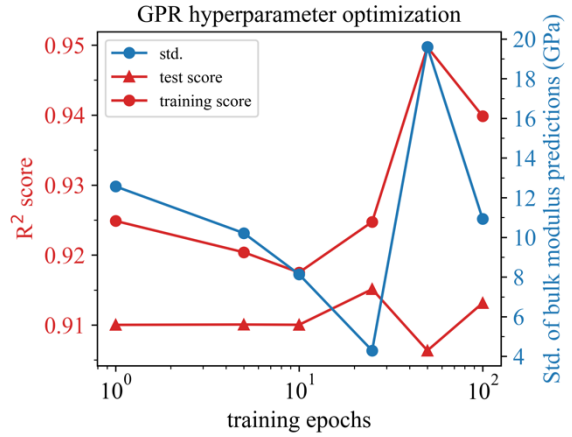

Figure S86: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.1, descriptor set A, and bulk modulus as target property.

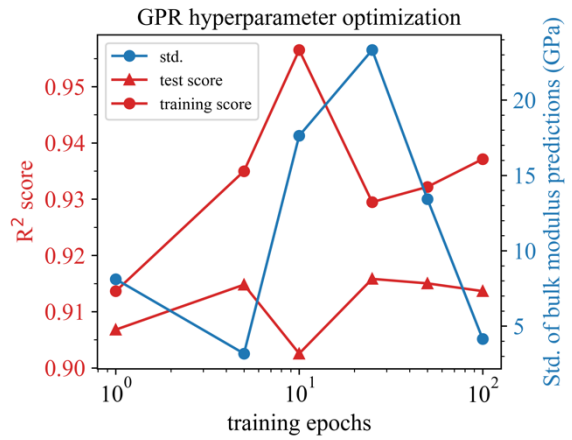

Figure S87: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 1.0, descriptor set A, and bulk modulus as target property.

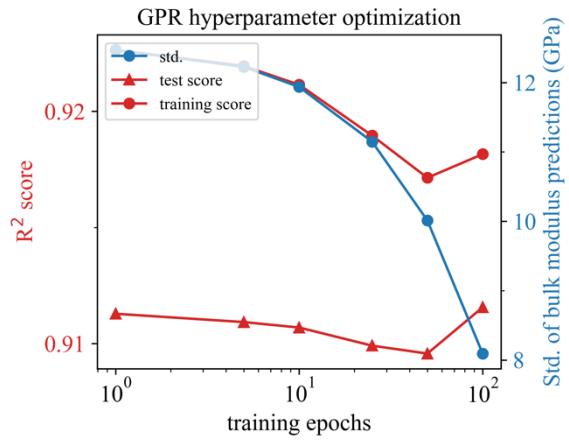

Figure S88: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.01, descriptor set C, and bulk modulus as target property.

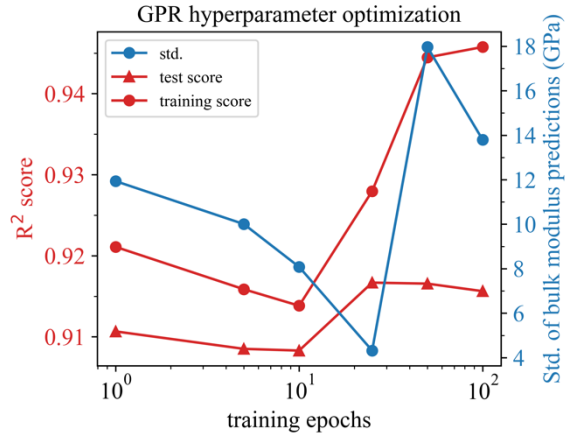

Figure S89: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.1, descriptor set C, and bulk modulus as target property.

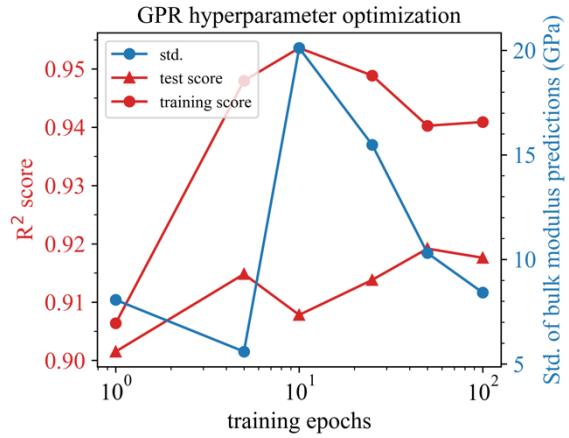

Figure S90: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 1.0, descriptor set C, and bulk modulus as target property.

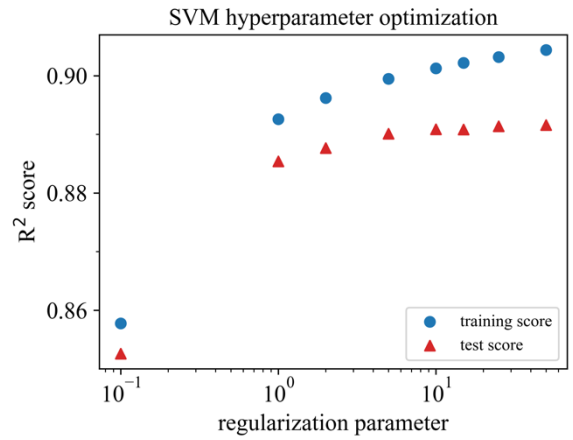

Figure S91: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.01$, descriptor set AC, and shear modulus as target property.

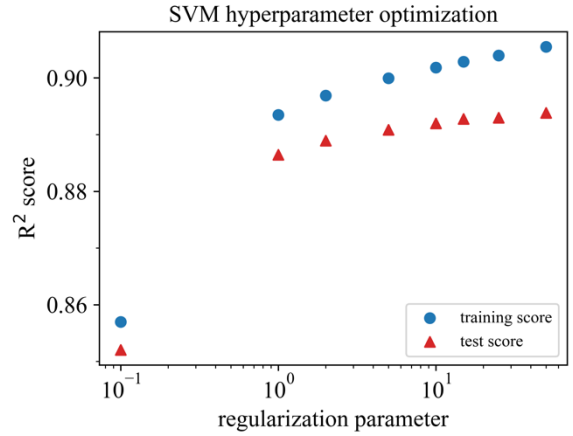

Figure S92: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.1$, descriptor set AC, and shear modulus as target property.

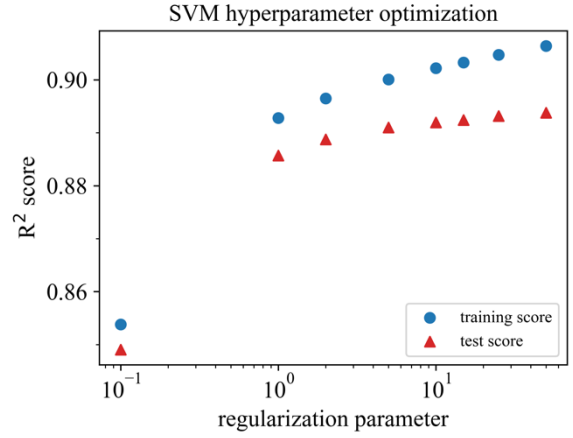

Figure S93: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.2$, descriptor set AC, and shear modulus as target property.

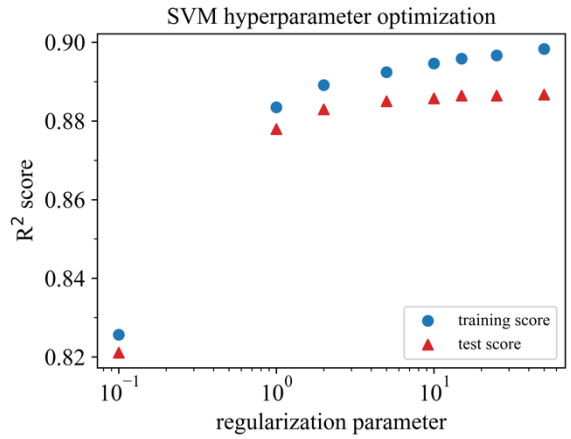

Figure S94: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.5$, descriptor set AC, and shear modulus as target property.

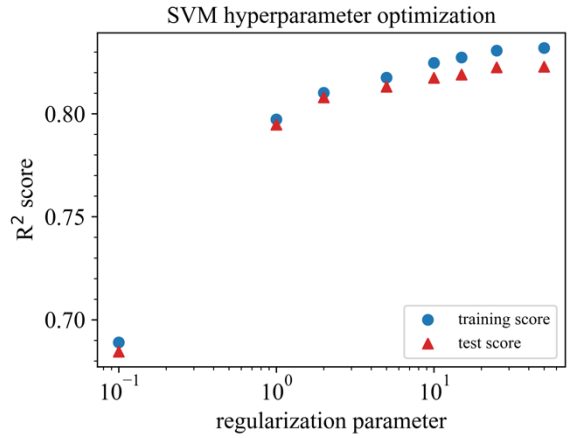

Figure S95: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 1.0$, descriptor set AC, and shear modulus as target property.

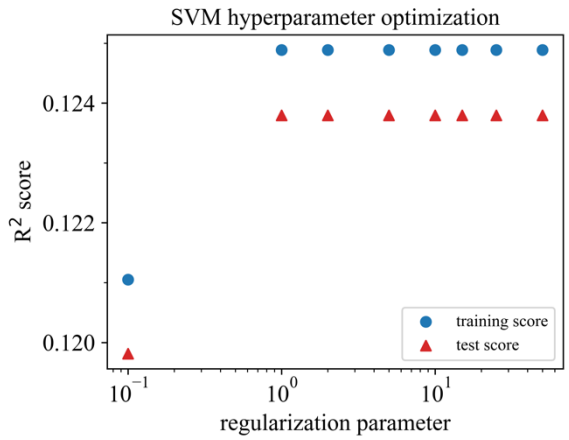

Figure S96: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 2.0$, descriptor set AC, and shear modulus as target property.

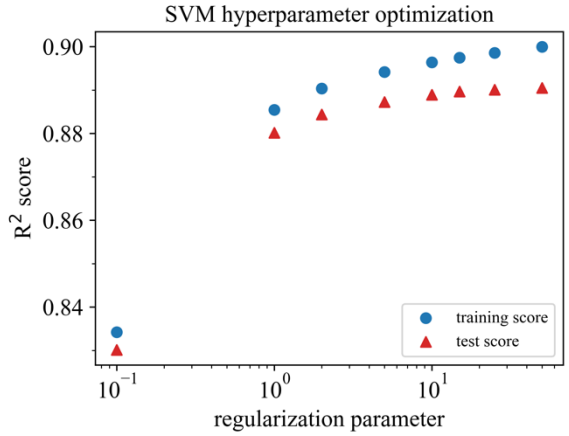

Figure S97: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.01$, descriptor set A, and shear modulus as target property.

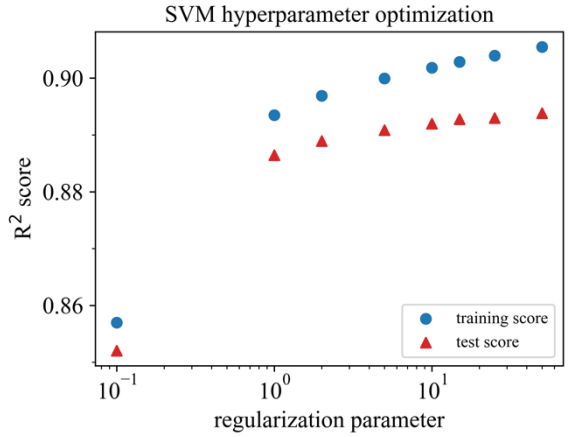

Figure S98: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.1$, descriptor set A, and shear modulus as target property.

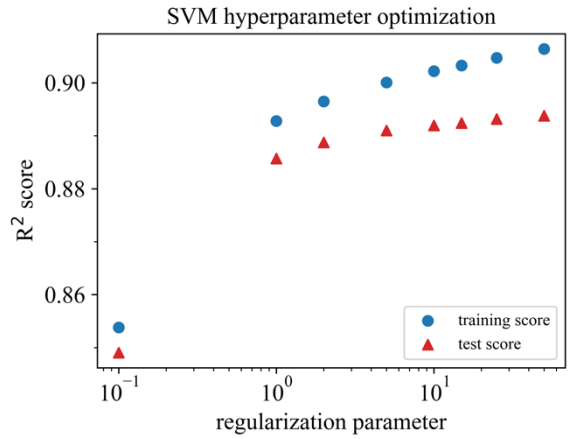

Figure S99: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.2$, descriptor set A, and shear modulus as target property.

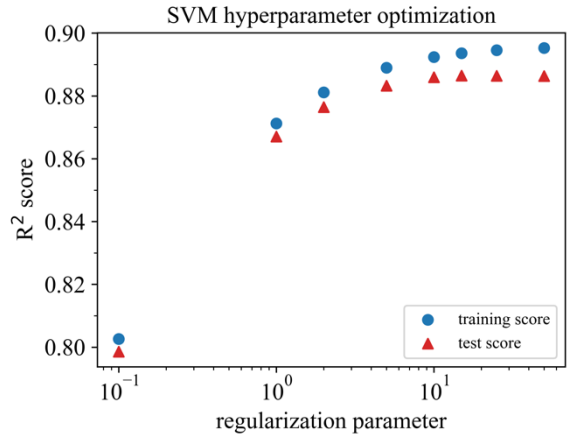

Figure S100: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.5$, descriptor set A, and shear modulus as target property.

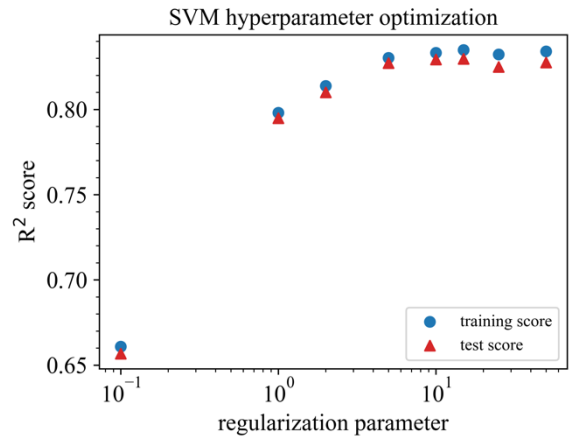

Figure S101: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, ϵ = 1.0, descriptor set A, and shear modulus as target property.

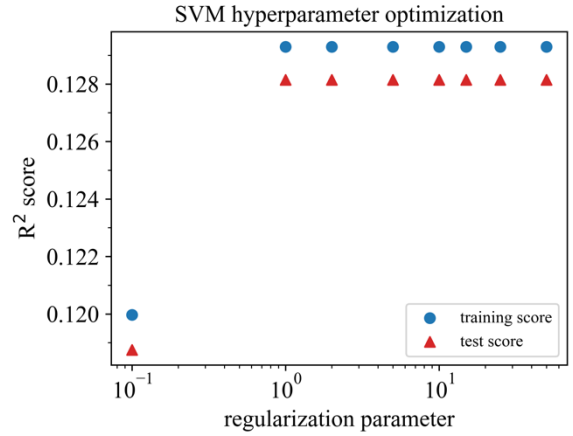

Figure S102: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, ϵ = 2.0, descriptor set A, and shear modulus as target property.

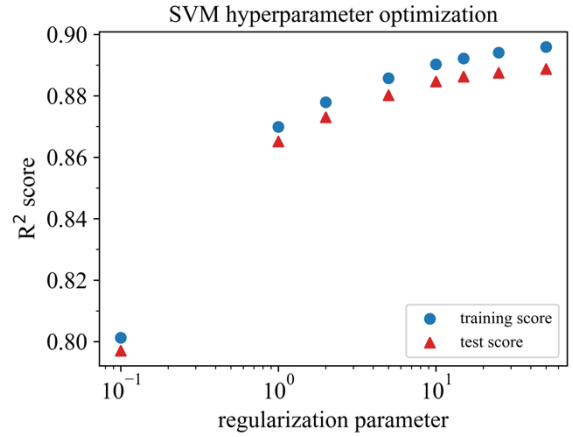

Figure S103: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, ϵ = 0.01, descriptor set C, and shear modulus as target property.

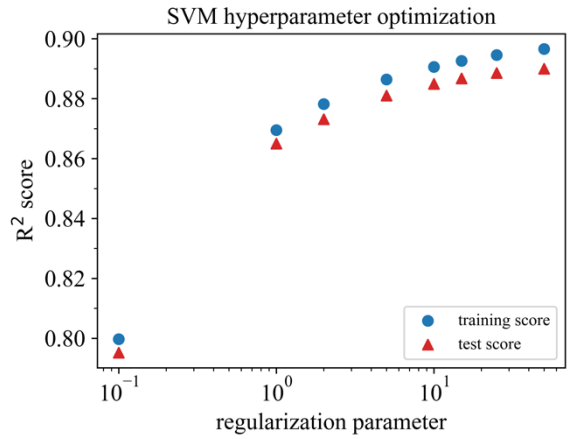

Figure S104: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.1$, descriptor set C, and shear modulus as target property.

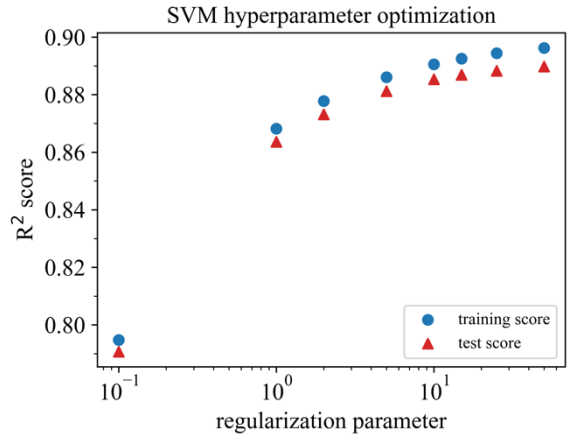

Figure S105: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.2$, descriptor set C, and shear modulus as target property.

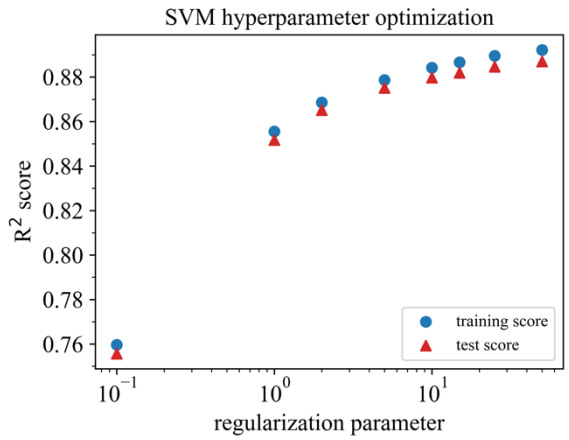

Figure S106: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.5$, descriptor set C, and shear modulus as target property.

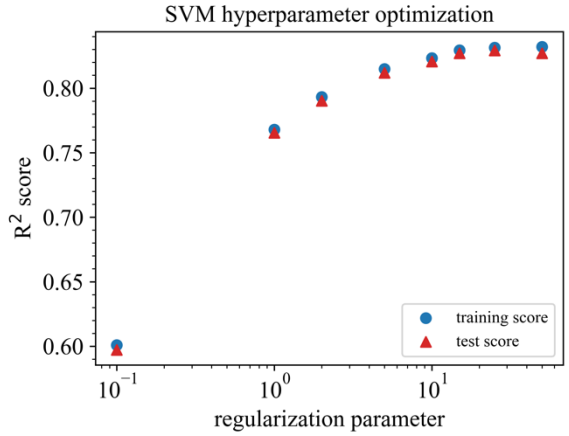

Figure S107: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 1.0$, descriptor set C, and shear modulus as target property.

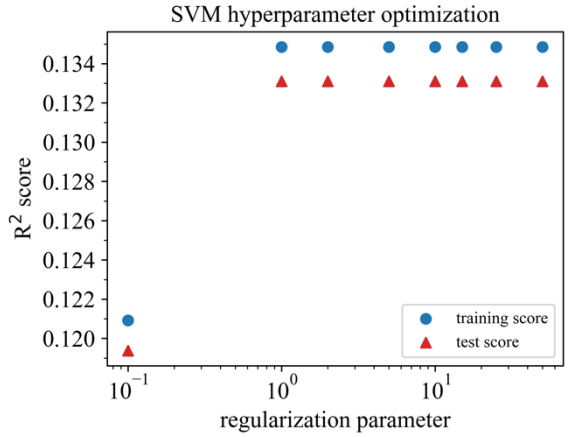

Figure S108: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 2.0$, descriptor set C, and shear modulus as target property.

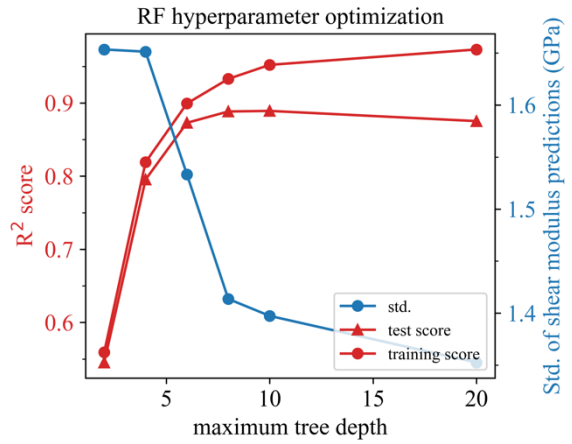

Figure S109: Hyperparameter optimization for RF trained with the V18 dataset, 4 estimators, descriptor set AC, and shear modulus as target property.

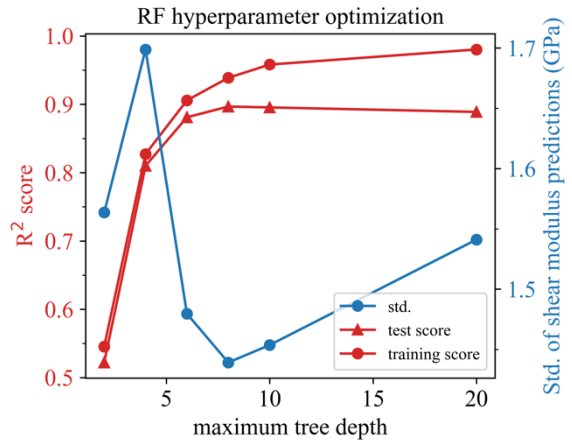

Figure S110: Hyperparameter optimization for RF trained with the V18 dataset, 8 estimators, descriptor set AC, and shear modulus as target property.

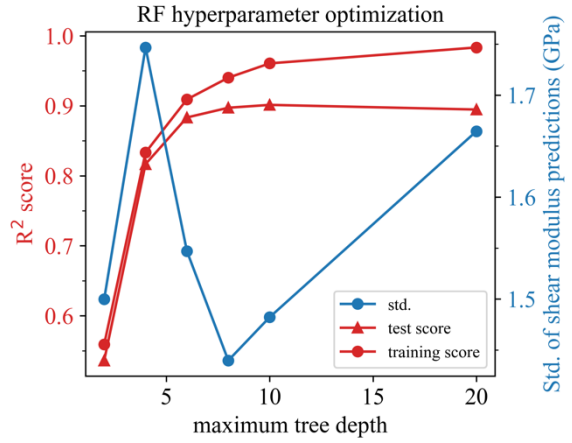

Figure S111: Hyperparameter optimization for RF trained with the V18 dataset, 16 estimators, descriptor set AC, and shear modulus as target property.

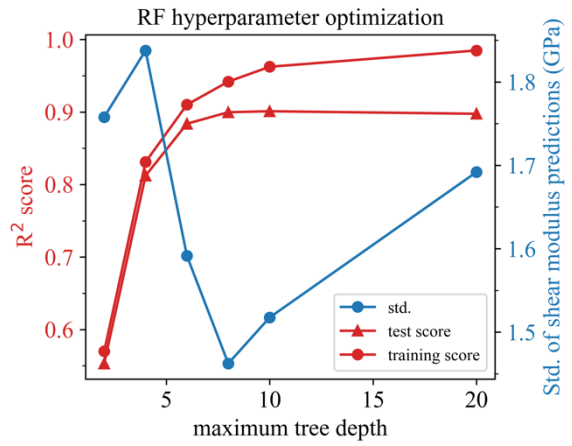

Figure S112: Hyperparameter optimization for RF trained with the V18 dataset, 32 estimators, descriptor set AC, and shear modulus as target property.

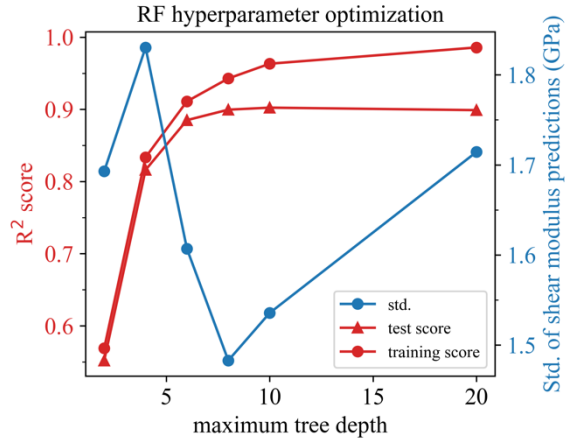

Figure S113: Hyperparameter optimization for RF trained with the V18 dataset, 64 estimators, descriptor set AC, and shear modulus as target property.

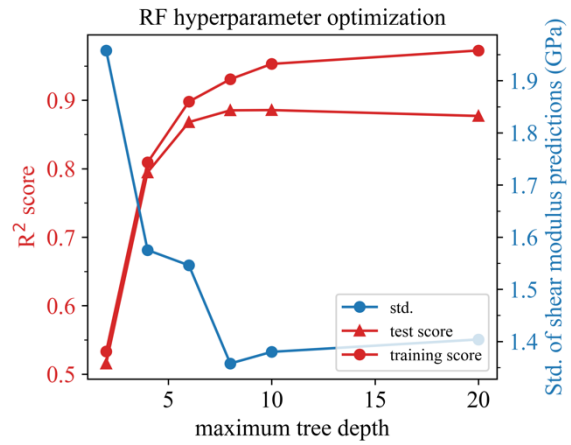

Figure S114: Hyperparameter optimization for RF trained with the V18 dataset, 4 estimators, descriptor set A, and shear modulus as target property.

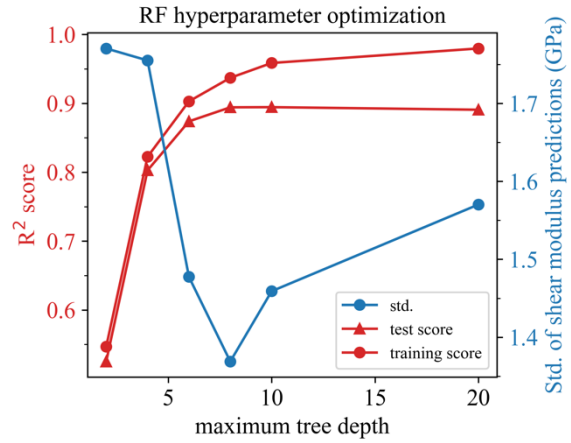

Figure S115: Hyperparameter optimization for RF trained with the V18 dataset, 8 estimators, descriptor set A, and shear modulus as target property.

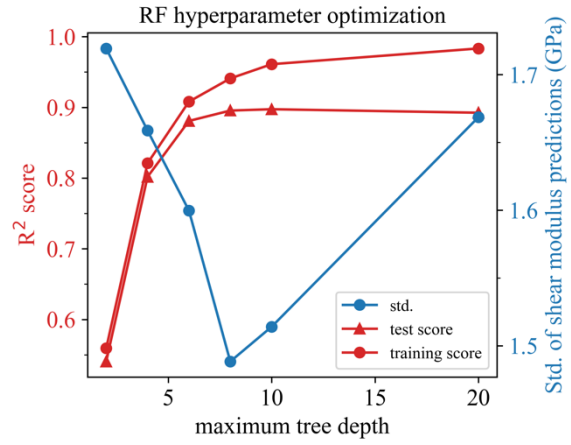

Figure S116: Hyperparameter optimization for RF trained with the V18 dataset, 16 estimators, descriptor set A, and shear modulus as target property.

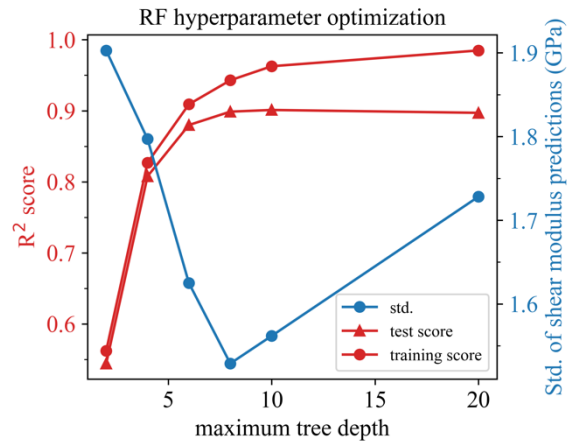

Figure S117: Hyperparameter optimization for RF trained with the V18 dataset, 32 estimators, descriptor set A, and shear modulus as target property.

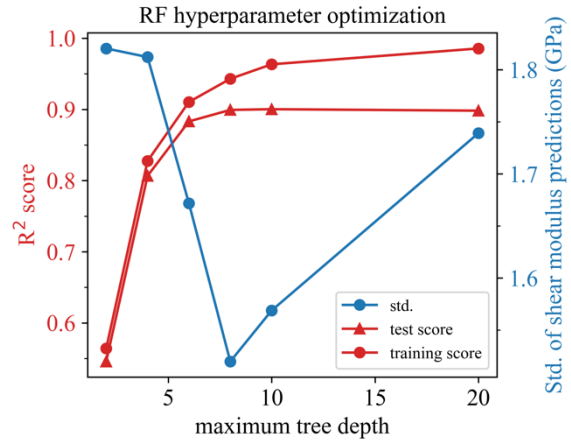

Figure S118: Hyperparameter optimization for RF trained with the V18 dataset, 64 estimators, descriptor set A, and shear modulus as target property.

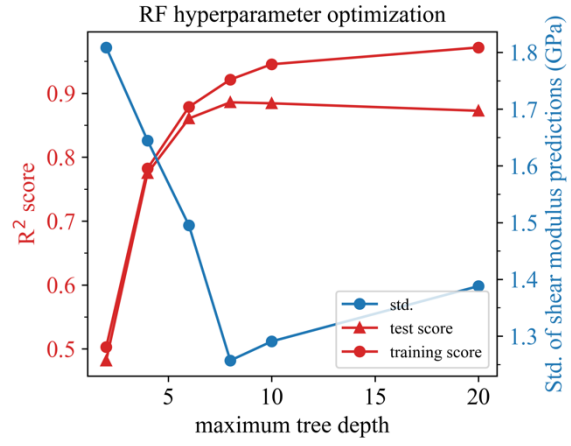

Figure S119: Hyperparameter optimization for RF trained with the V18 dataset, 4 estimators, descriptor set C, and shear modulus as target property.

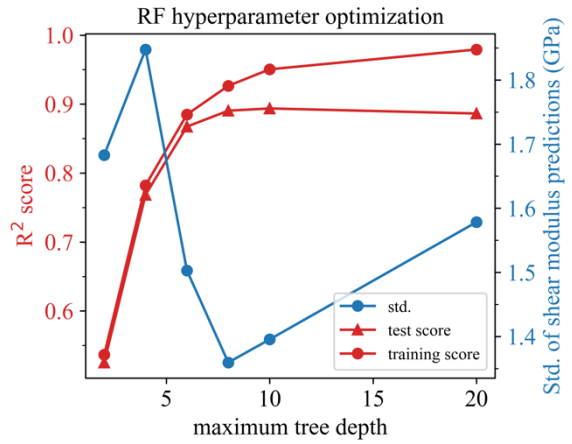

Figure S120: Hyperparameter optimization for RF trained with the V18 dataset, 8 estimators, descriptor set C, and shear modulus as target property.

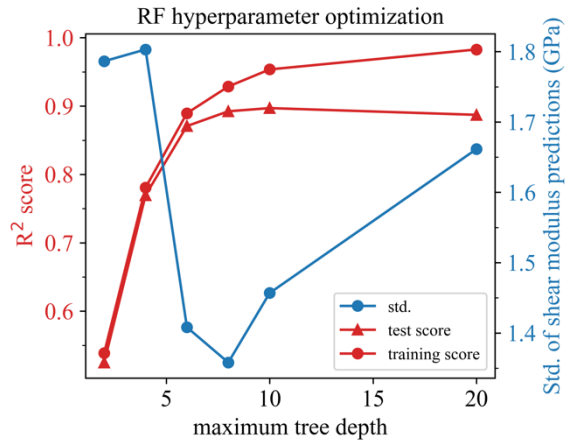

Figure S121: Hyperparameter optimization for RF trained with the V18 dataset, 16 estimators, descriptor set C, and shear modulus as target property.

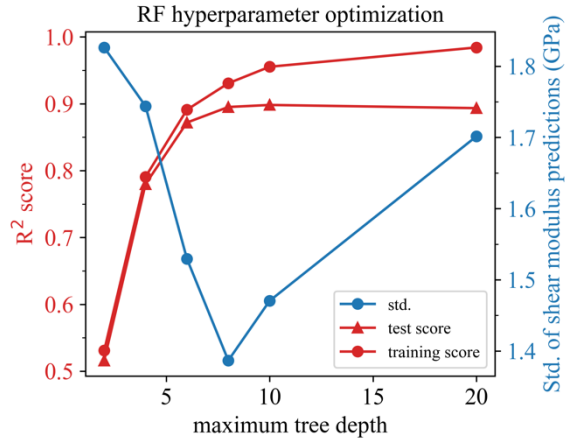

Figure S122: Hyperparameter optimization for RF trained with the V18 dataset, 32 estimators, descriptor set C, and shear modulus as target property.

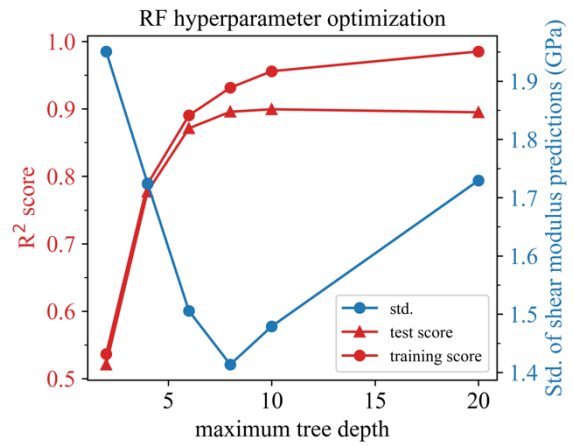

Figure S123: Hyperparameter optimization for RF trained with the V18 dataset, 64 estimators, descriptor set C, and shear modulus as target property.

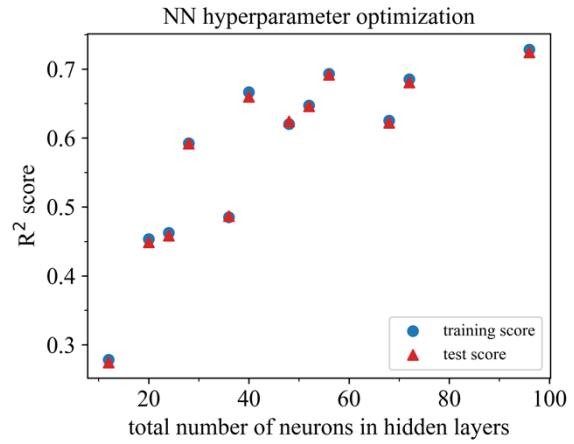

Figure S124: Hyperparameter optimization for NN trained with the V18 dataset, descriptor set AC, and shear modulus as target property.

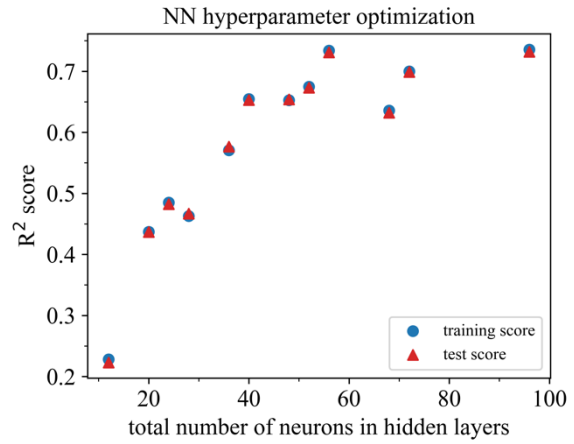

Figure S125: Hyperparameter optimization for NN trained with the V18 dataset, descriptor set A, and shear modulus as target property.

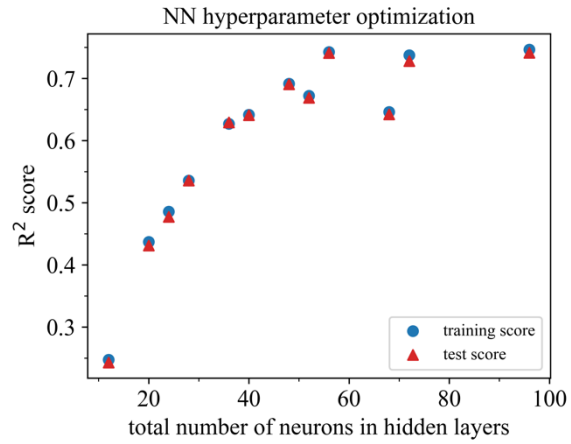

Figure S126: Hyperparameter optimization for NN trained with the V18 dataset, descriptor set C, and shear modulus as target property.

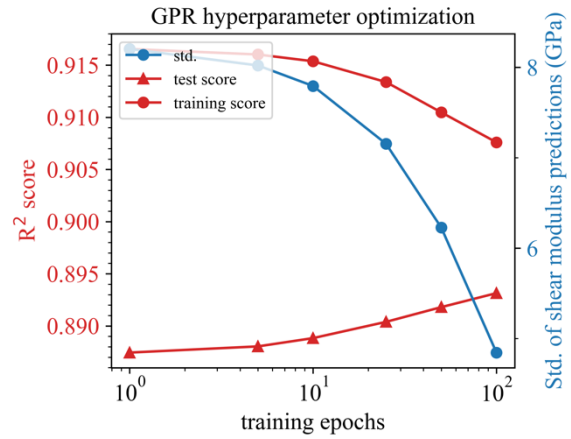

Figure S127: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.01, descriptor set AC, and shear modulus as target property.

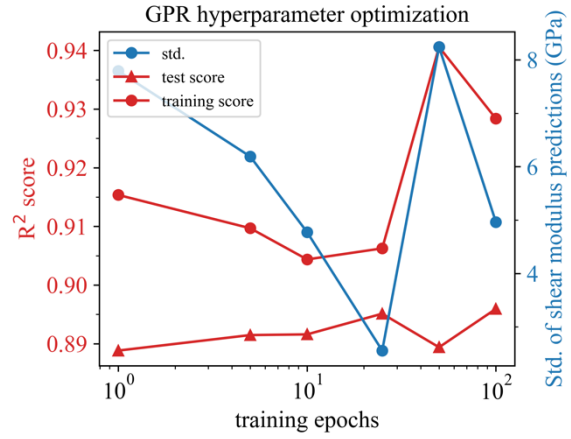

Figure S128: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.1, descriptor set AC, and shear modulus as target property.

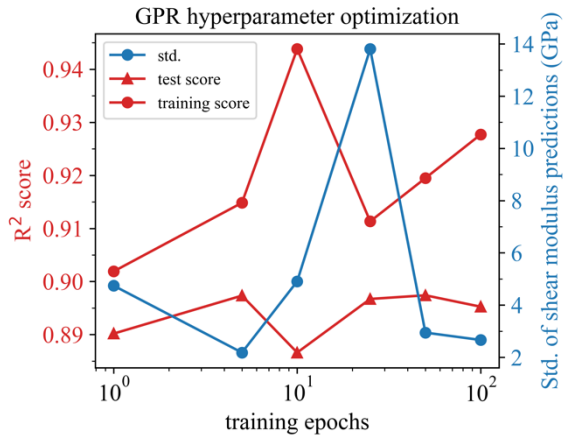

Figure S129: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 1.0, descriptor set AC, and shear modulus as target property.

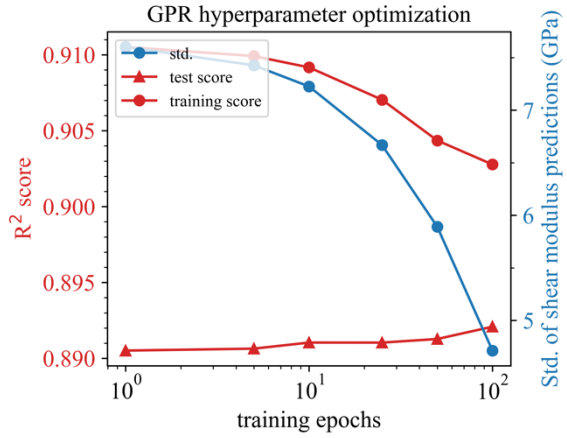

Figure S130: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.01, descriptor set A, and shear modulus as target property.

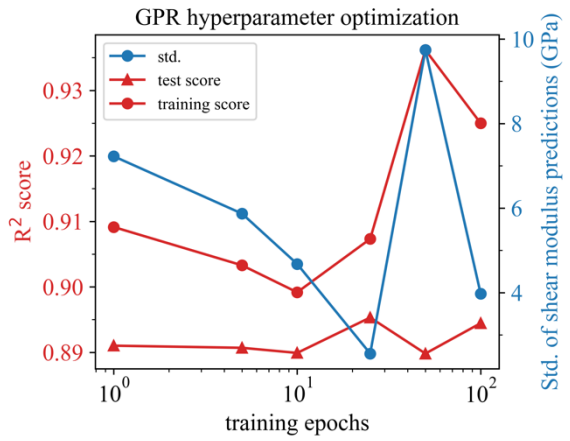

Figure S131: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.1, descriptor set A, and shear modulus as target property.

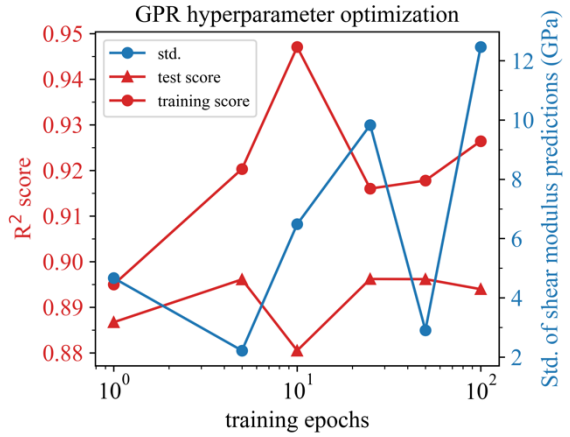

Figure S132: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 1.0, descriptor set A, and shear modulus as target property.

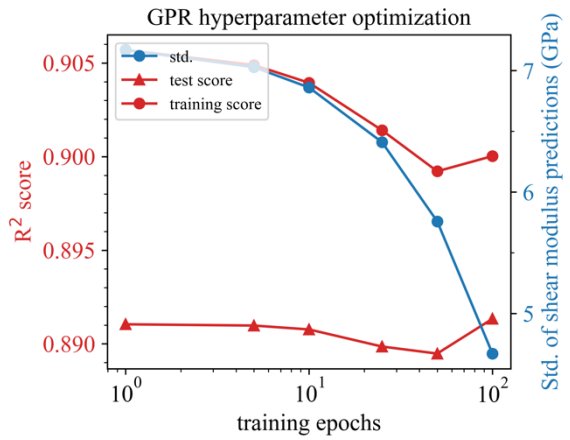

Figure S133: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.01, descriptor set C, and shear modulus as target property.

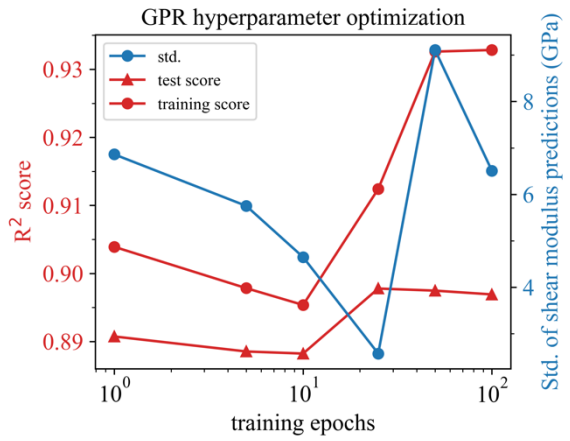

Figure S134: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.1, descriptor set C, and shear modulus as target property.

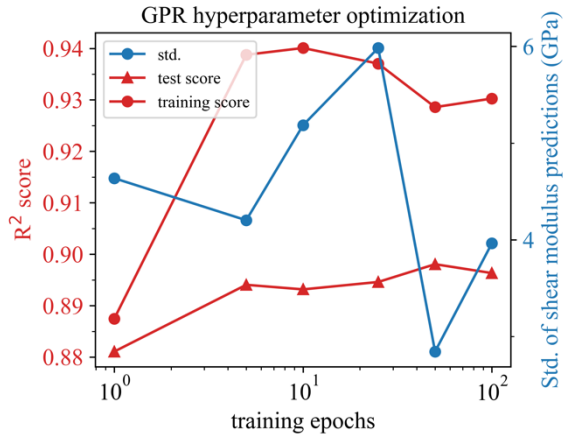

Figure S135: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 1.0, descriptor set C, and shear modulus as target property.

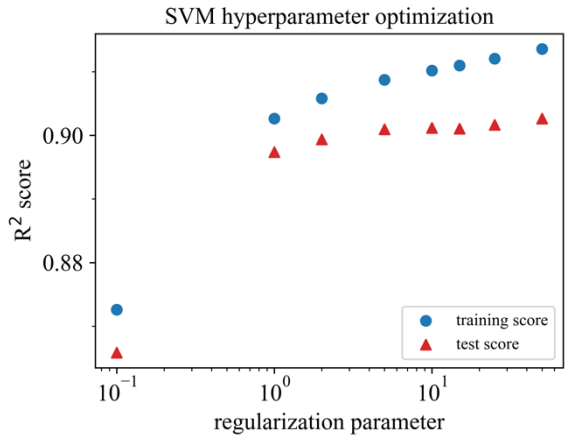

Figure S136: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.01$, descriptor set AC, and Young's modulus as target property.

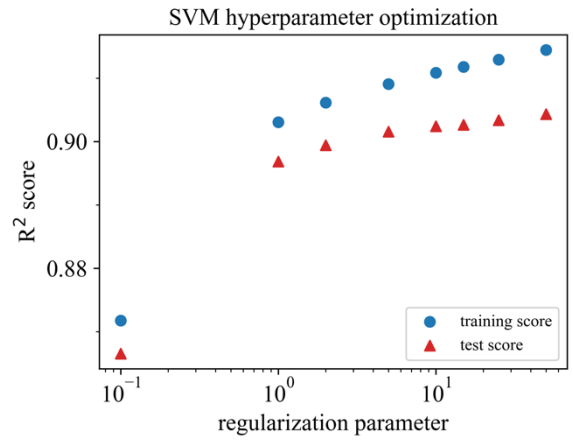

Figure S137: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.1$, descriptor set AC, and Young's modulus as target property.

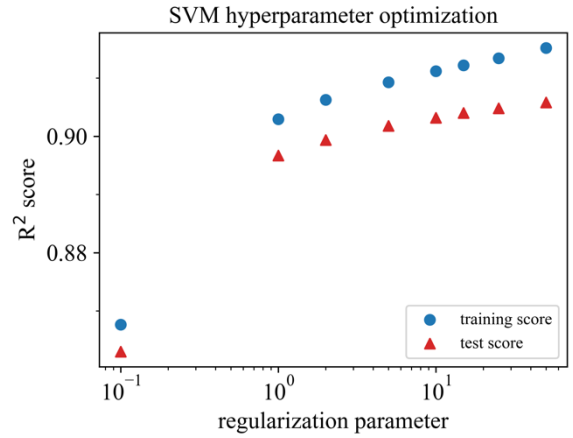

Figure S138: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.2$, descriptor set AC, and Young's modulus as target property.

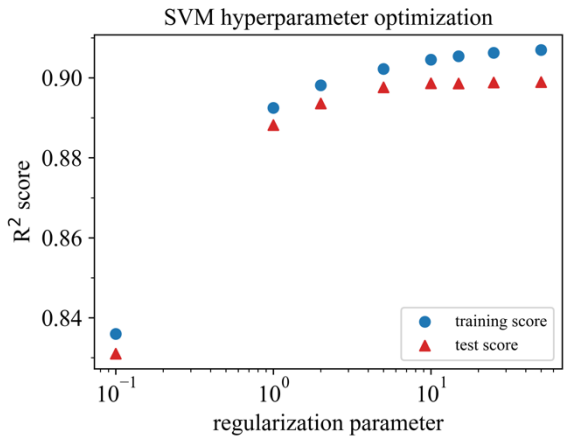

Figure S139: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, ϵ = 0.5, descriptor set AC, and Young's modulus as target property.

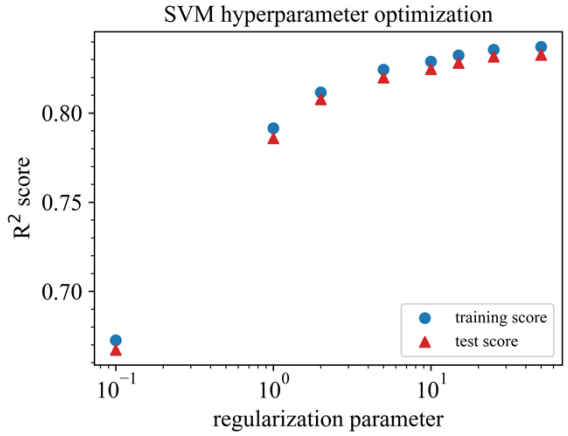

Figure S140: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, ϵ = 1.0, descriptor set AC, and Young's modulus as target property.

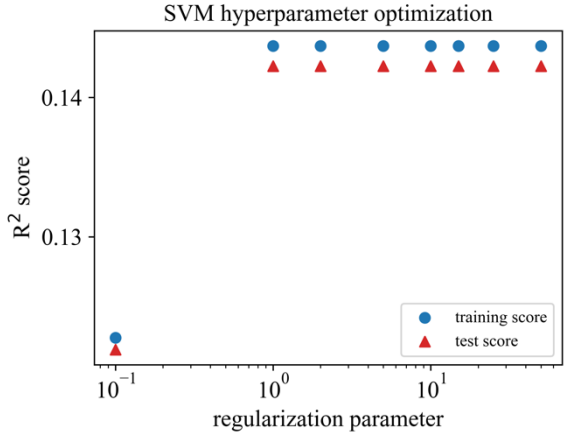

Figure S141: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, ϵ = 2.0, descriptor set AC, and Young's modulus as target property.

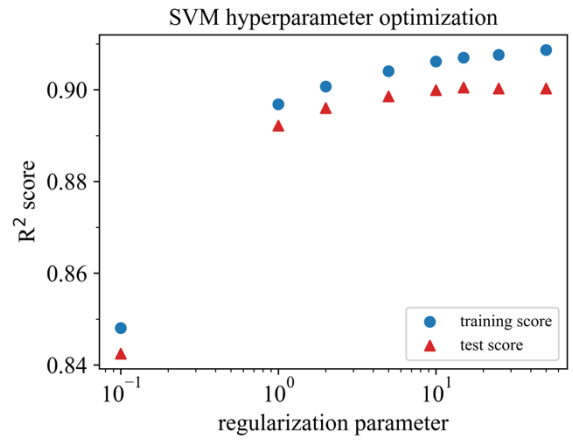

Figure S142: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.01$, descriptor set A, and Young's modulus as target property.

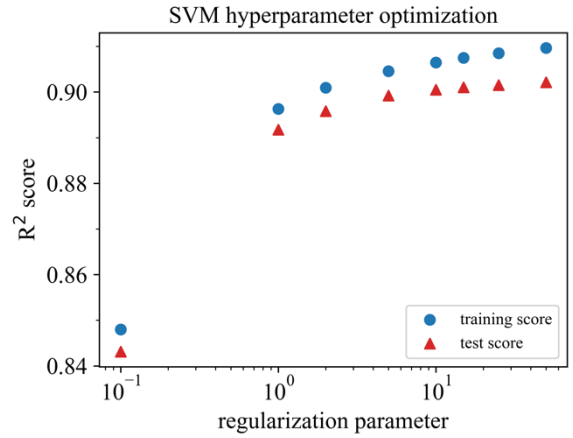

Figure S143: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.1$, descriptor set A, and Young's modulus as target property.

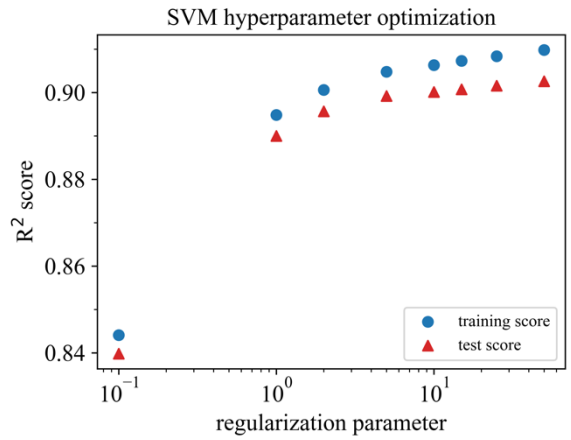

Figure S144: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.2$, descriptor set A, and Young's modulus as target property.

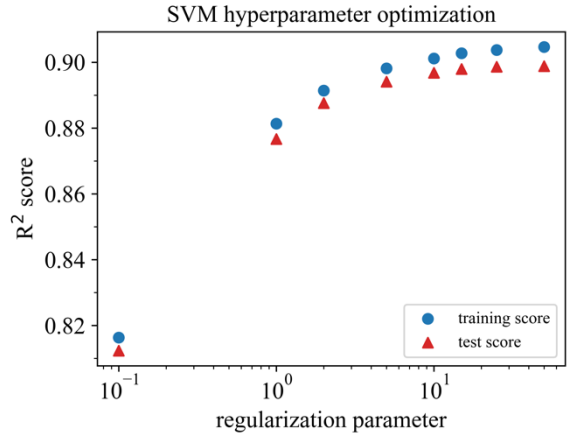

Figure S145: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.5$, descriptor set A, and Young's modulus as target property.

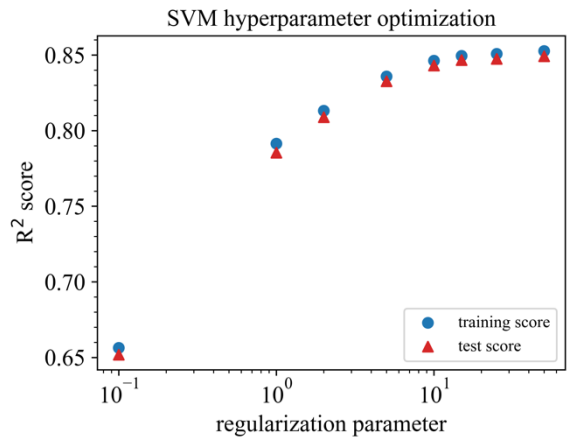

Figure S146: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 1.0$, descriptor set A, and Young's modulus as target property.

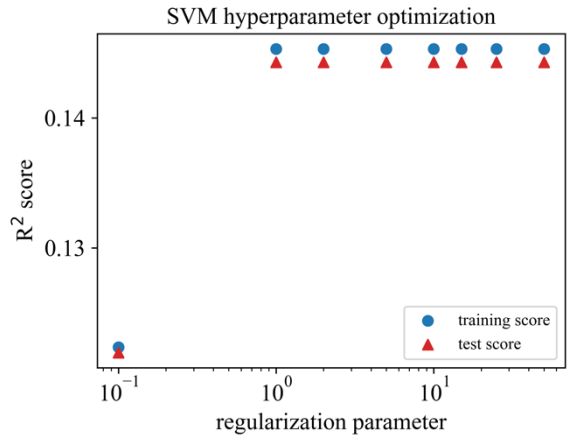

Figure S147: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 2.0$, descriptor set A, and Young's modulus as target property.

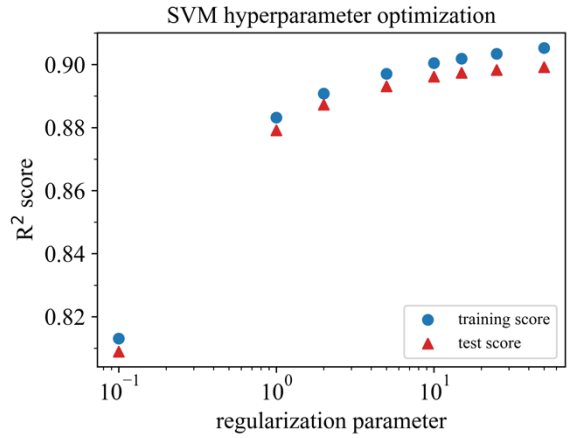

Figure S148: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.01$, descriptor set C, and Young's modulus as target property.

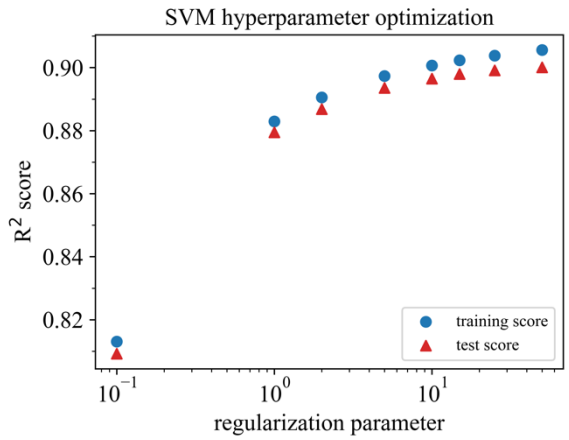

Figure S149: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.1$, descriptor set C, and Young's modulus as target property.

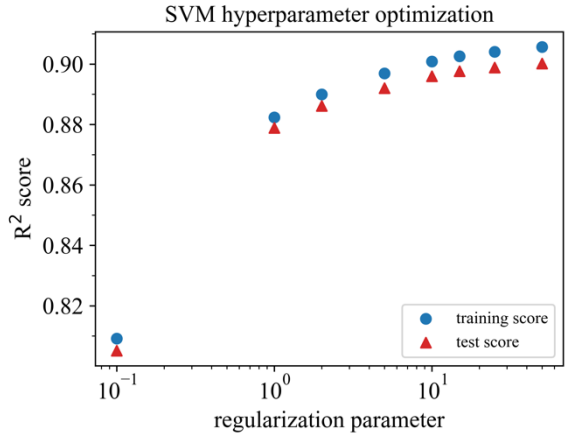

Figure S150: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.2$, descriptor set C, and Young's modulus as target property.

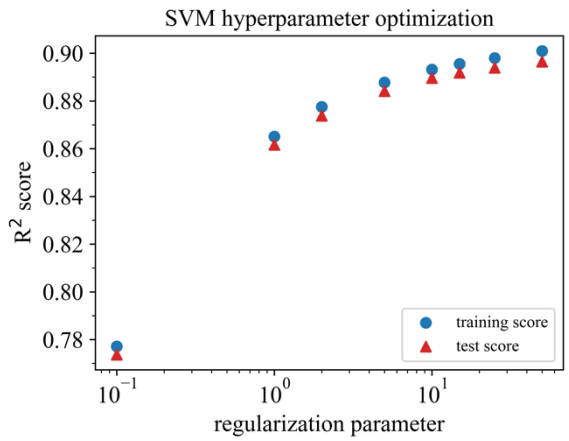

Figure S151: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, $\epsilon = 0.5$, descriptor set C, and Young's modulus as target property.

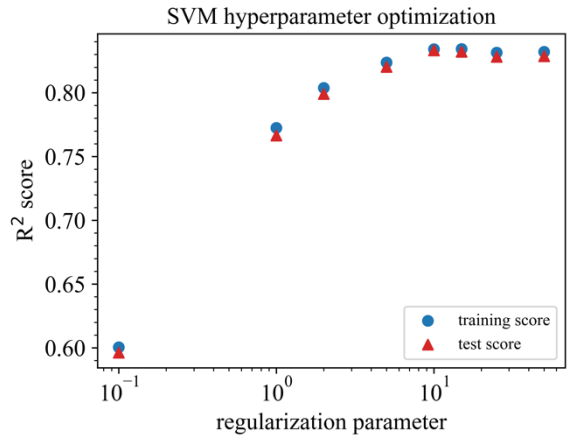

Figure S152: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, ϵ = 1.0, descriptor set C, and Young's modulus as target property.

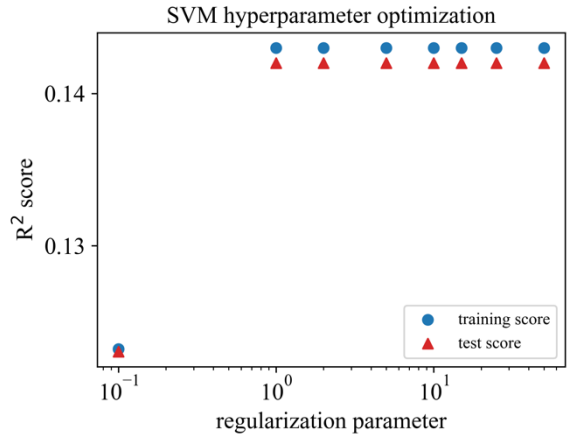

Figure S153: Hyperparameter optimization for SVM (SVR) trained with the V18 dataset, ϵ = 2.0, descriptor set C, and Young's modulus as target property.

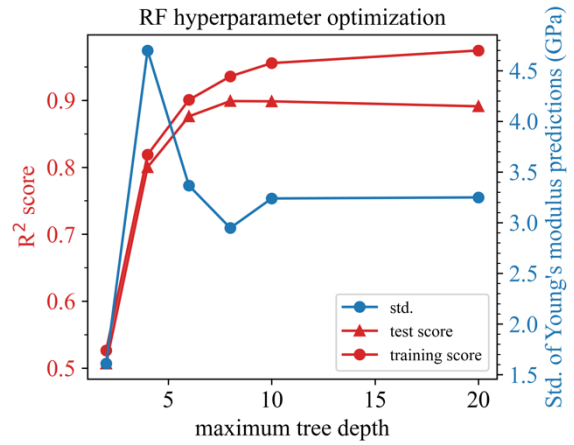

Figure S154: Hyperparameter optimization for RF trained with the V18 dataset, 4 estimators, descriptor set AC, and Young's modulus as target property.

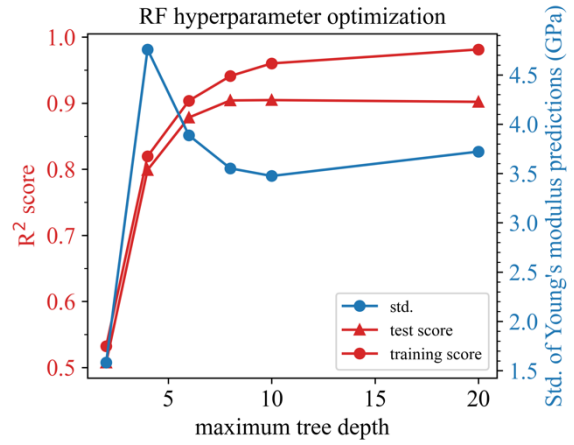

Figure S155: Hyperparameter optimization for RF trained with the V18 dataset, 8 estimators, descriptor set AC, and Young's modulus as target property.

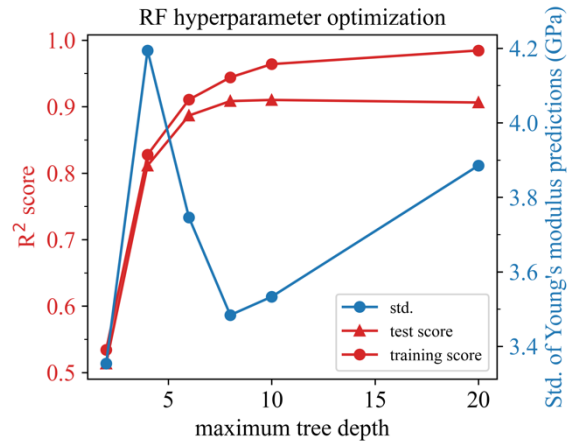

Figure S156: Hyperparameter optimization for RF trained with the V18 dataset, 16 estimators, descriptor set AC, and Young's modulus as target property.

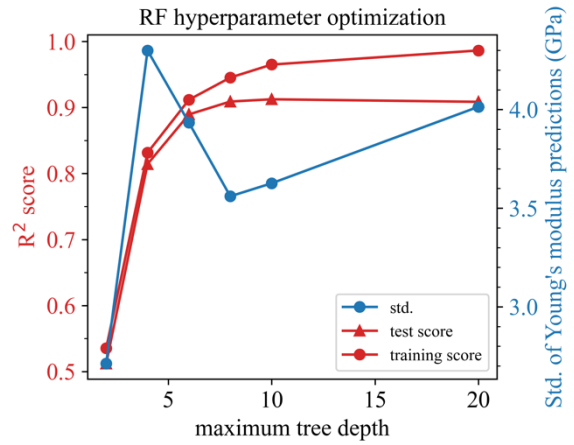

Figure S157: Hyperparameter optimization for RF trained with the V18 dataset, 32 estimators, descriptor set AC, and Young's modulus as target property.

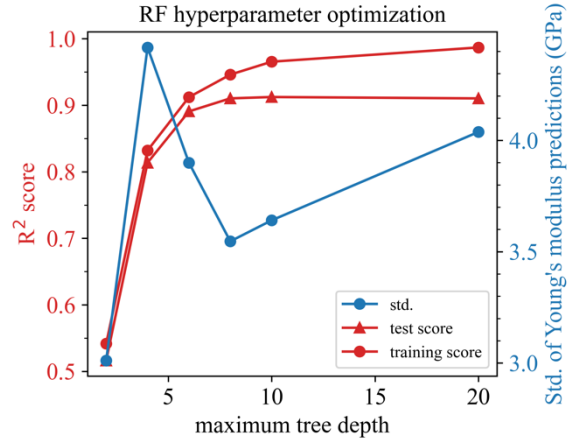

Figure S158: Hyperparameter optimization for RF trained with the V18 dataset, 64 estimators, descriptor set AC, and Young's modulus as target property.

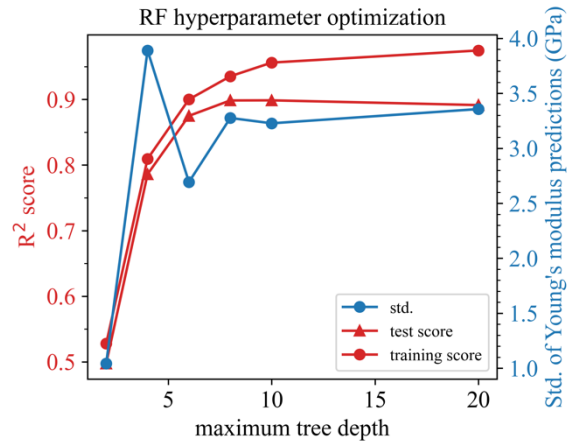

Figure S159: Hyperparameter optimization for RF trained with the V18 dataset, 4 estimators, descriptor set A, and Young's modulus as target property.

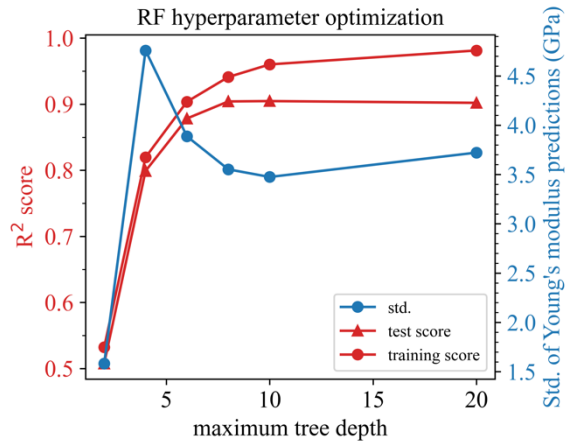

Figure S160: Hyperparameter optimization for RF trained with the V18 dataset, 8 estimators, descriptor set A, and Young's modulus as target property.

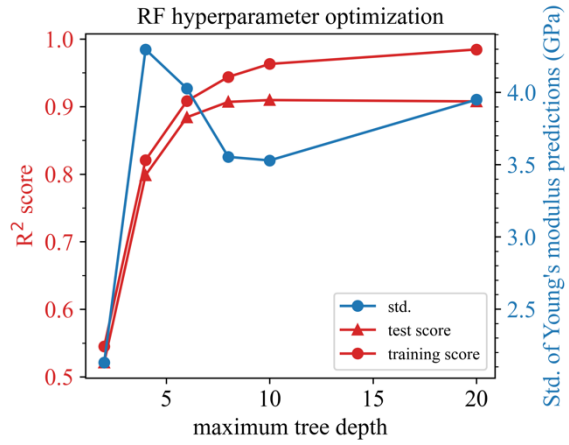

Figure S161: Hyperparameter optimization for RF trained with the V18 dataset, 16 estimators, descriptor set A, and Young's modulus as target property.

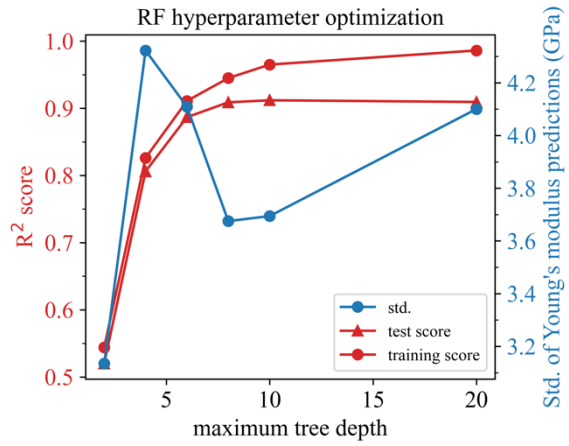

Figure S162: Hyperparameter optimization for RF trained with the V18 dataset, 32 estimators, descriptor set A, and Young's modulus as target property.

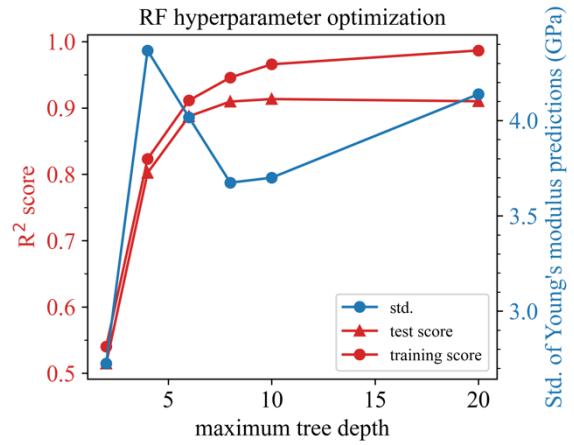

Figure S163: Hyperparameter optimization for RF trained with the V18 dataset, 64 estimators, descriptor set A, and Young's modulus as target property.

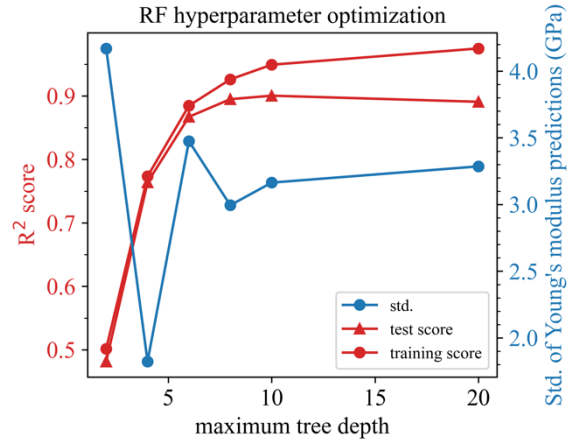

Figure S164: Hyperparameter optimization for RF trained with the V18 dataset, 4 estimators, descriptor set C, and Young's modulus as target property.

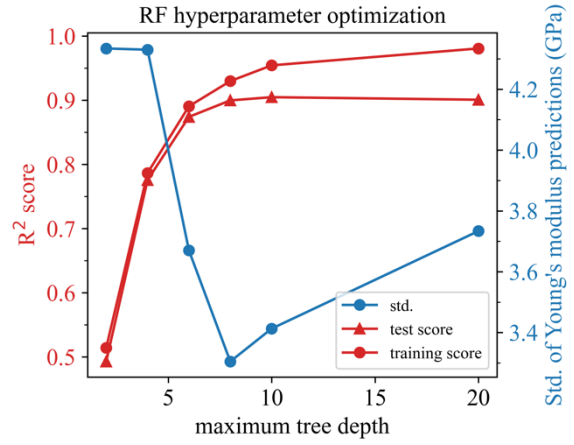

Figure S165: Hyperparameter optimization for RF trained with the V18 dataset, 8 estimators, descriptor set C, and Young's modulus as target property.

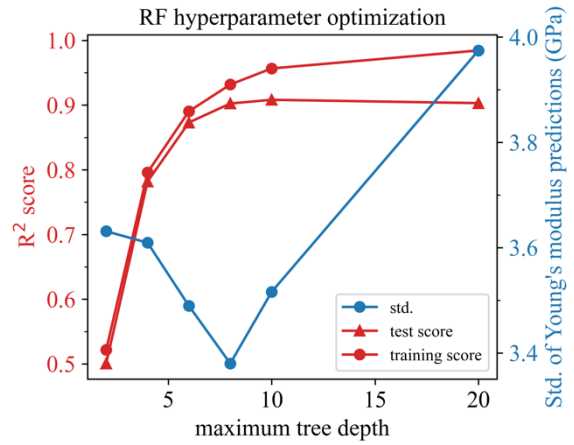

Figure S166: Hyperparameter optimization for RF trained with the V18 dataset, 16 estimators, descriptor set C, and Young's modulus as target property.

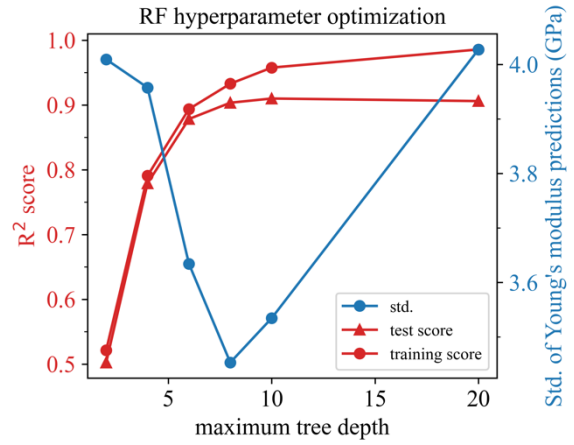

Figure S167: Hyperparameter optimization for RF trained with the V18 dataset, 32 estimators, descriptor set C, and Young's modulus as target property.

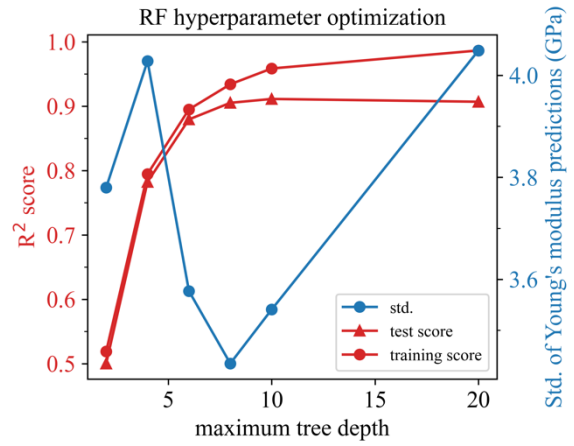

Figure S168: Hyperparameter optimization for RF trained with the V18 dataset, 64 estimators, descriptor set C, and Young's modulus as target property.

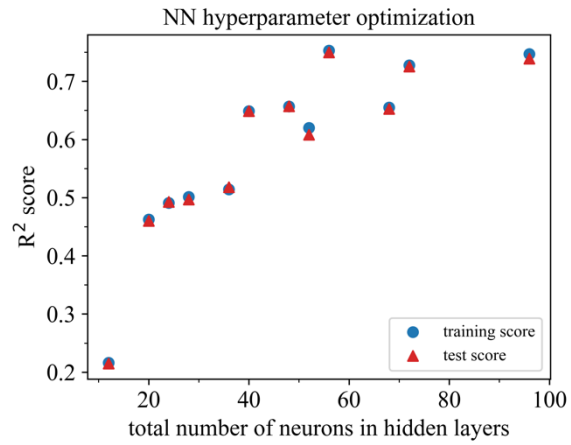

Figure S169: Hyperparameter optimization for NN trained with the V18 dataset, descriptor set AC, and Young's modulus as target property.

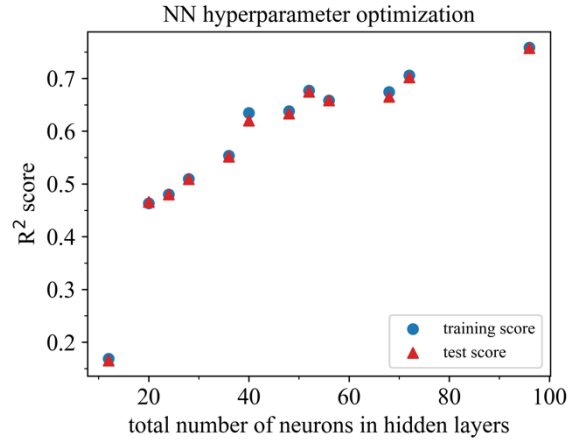

Figure S170: Hyperparameter optimization for NN trained with the V18 dataset, descriptor set A, and Young's modulus as target property.

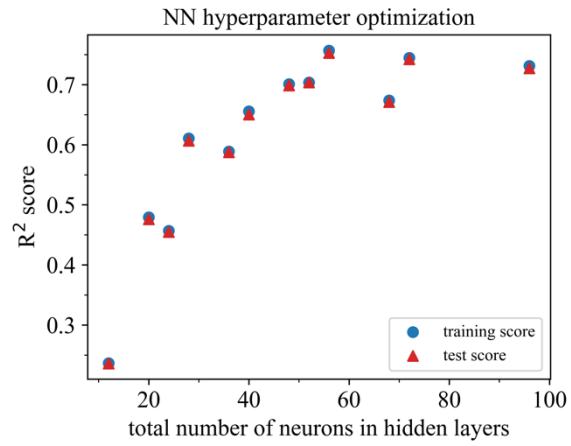

Figure S171: Hyperparameter optimization for NN trained with the V18 dataset, descriptor set C, and Young's modulus as target property.

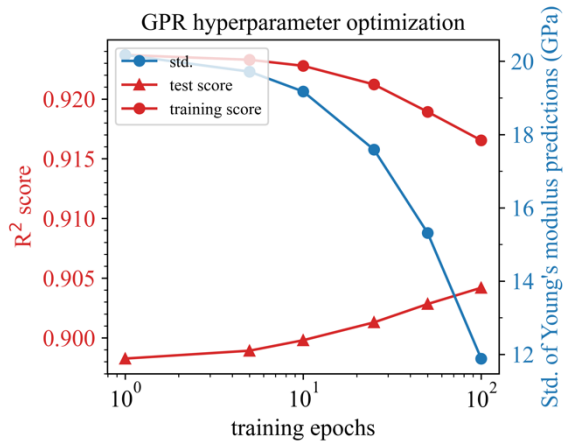

Figure S172: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.01, descriptor set AC, and Young's modulus as target property.

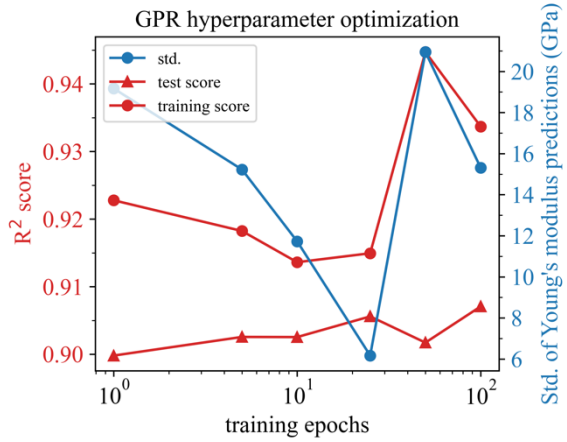

Figure S173: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.1, descriptor set AC, and Young's modulus as target property.

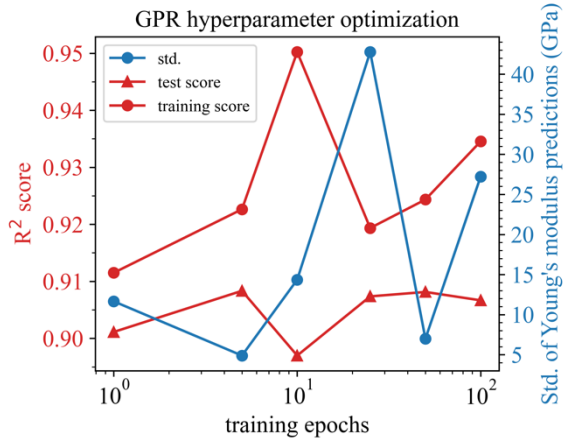

Figure S174: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 1.0, descriptor set AC, and Young's modulus as target property.

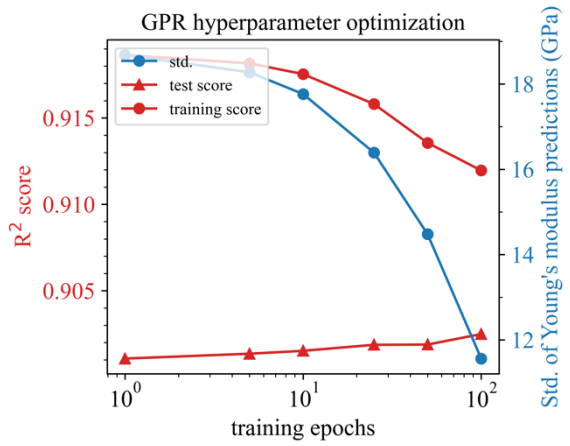

Figure S175: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.01, descriptor set A, and Young's modulus as target property.

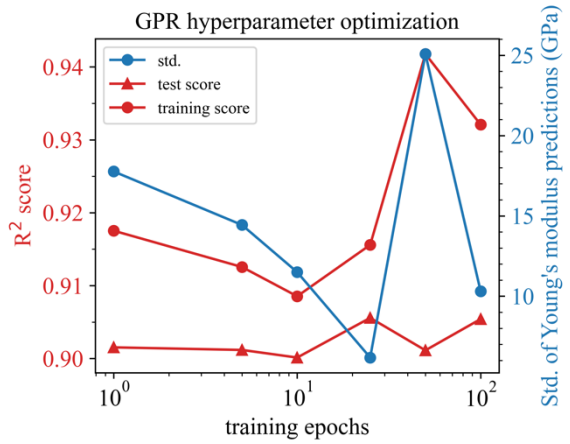

Figure S176: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.1, descriptor set A, and Young's modulus as target property.

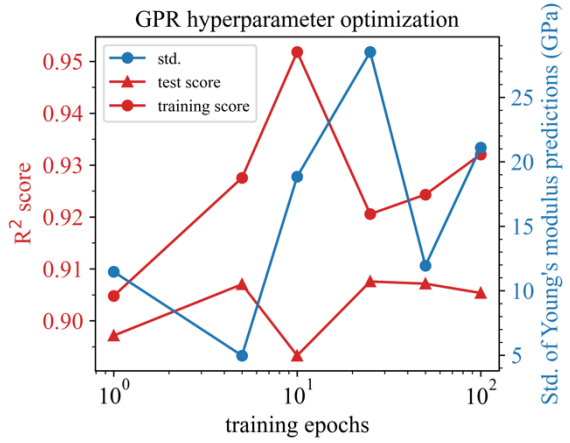

Figure S177: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 1.0, descriptor set A, and Young's modulus as target property.

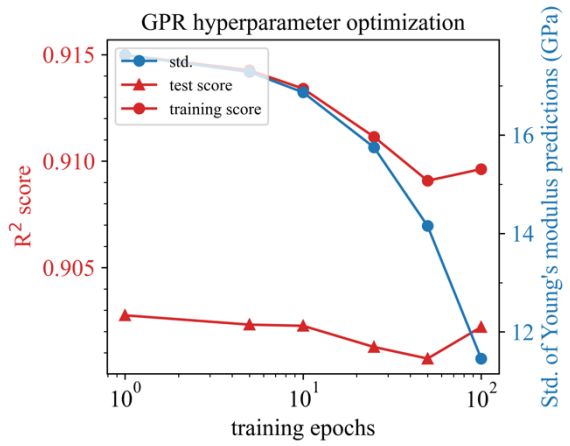

Figure S178: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.01, descriptor set C, and Young's modulus as target property.

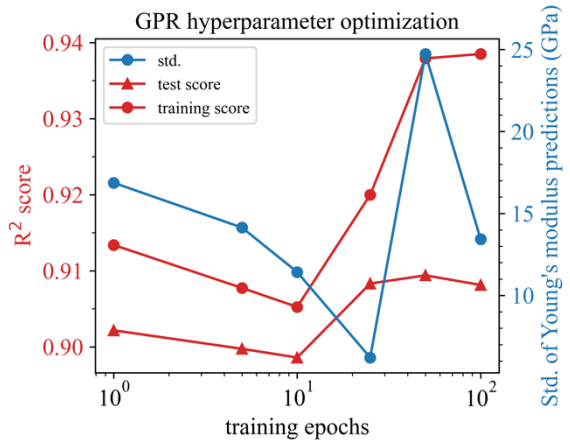

Figure S179: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 0.1, descriptor set C, and Young's modulus as target property.

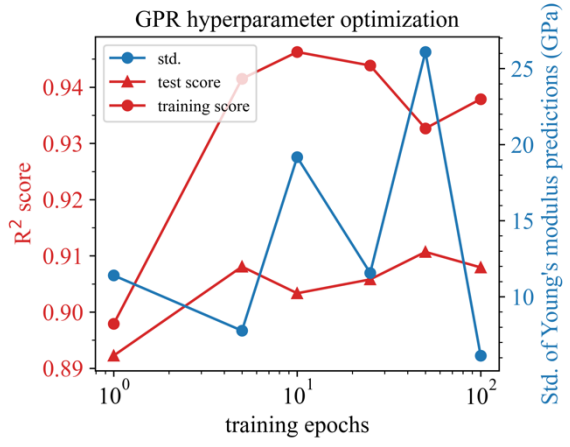

Figure S180: Hyperparameter optimization for GPR trained with the V18 dataset, a learning rate of 1.0, descriptor set C, and Young's modulus as target property.

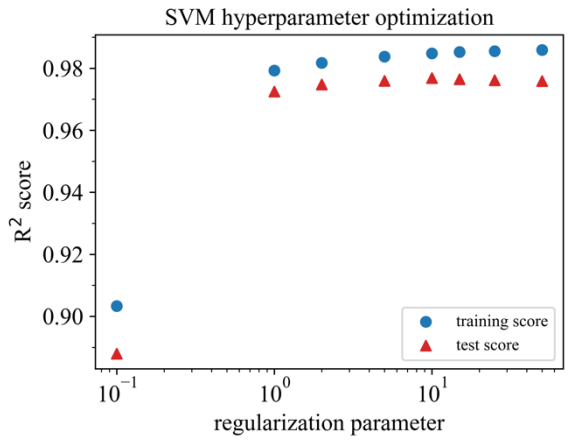

Figure S181: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.01$, descriptor set AC, and density as target property.

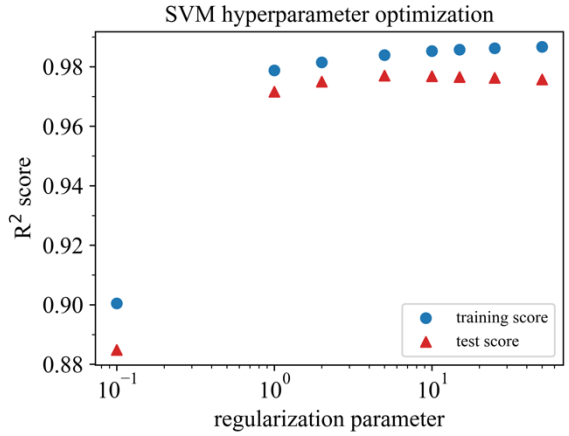

Figure S182: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.1$, descriptor set AC, and density as target property.

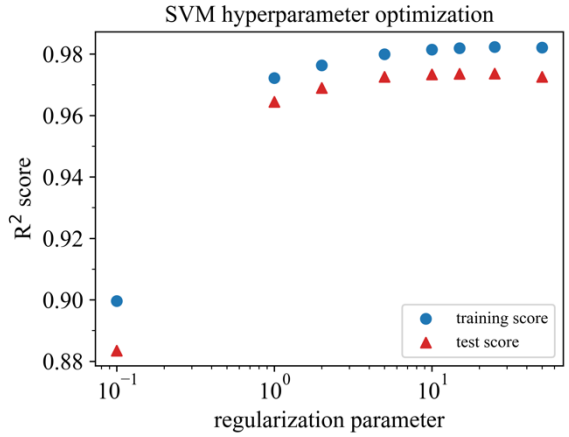

Figure S183: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.2$, descriptor set AC, and density as target property.

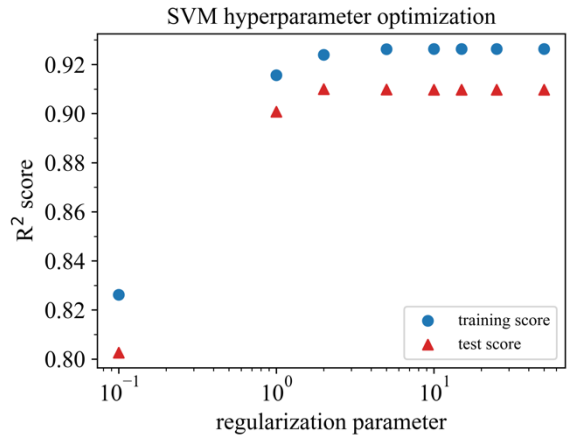

Figure S184: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.5$, descriptor set AC, and density as target property.

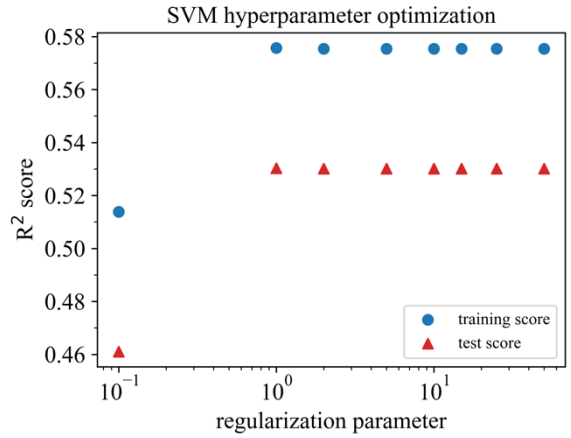

Figure S185: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 1.0$, descriptor set AC, and density as target property.

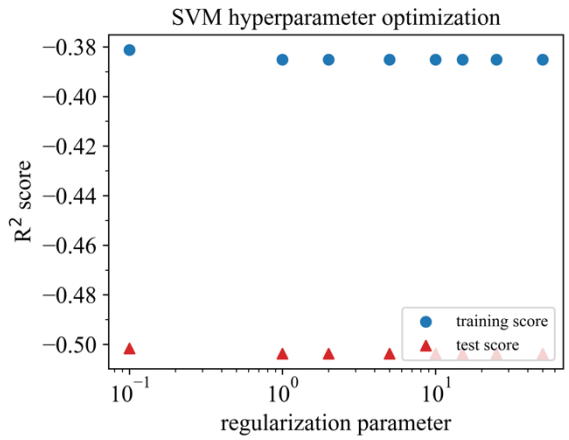

Figure S186: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 2.0$, descriptor set AC, and density as target property.

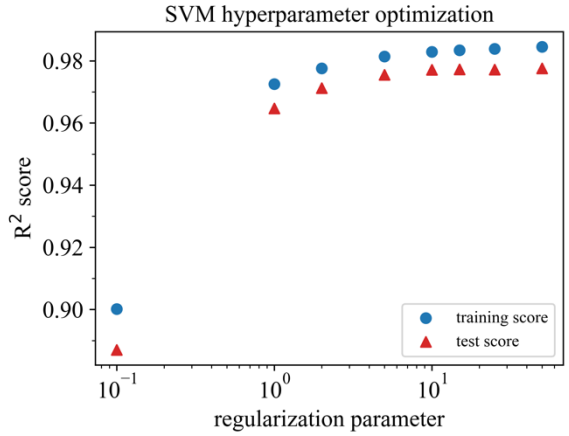

Figure S187: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.01$, descriptor set A, and density as target property.

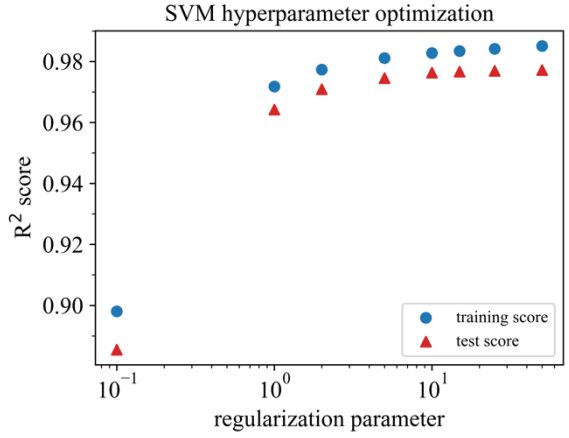

Figure S188: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.1$, descriptor set A, and density as target property.

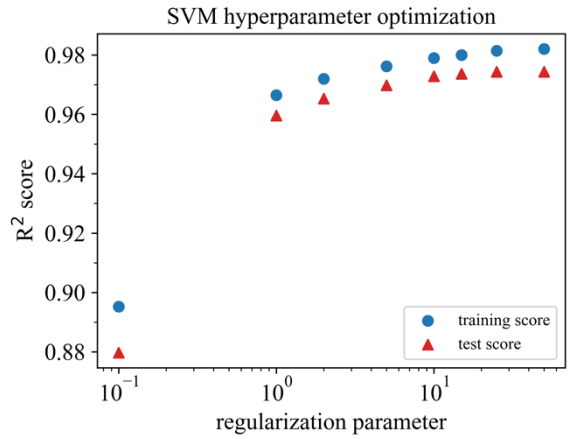

Figure S189: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.2$, descriptor set A, and density as target property.

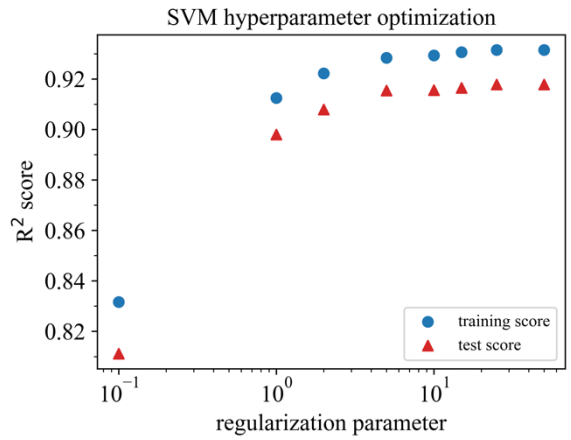

Figure S190: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.5$, descriptor set A, and density as target property.

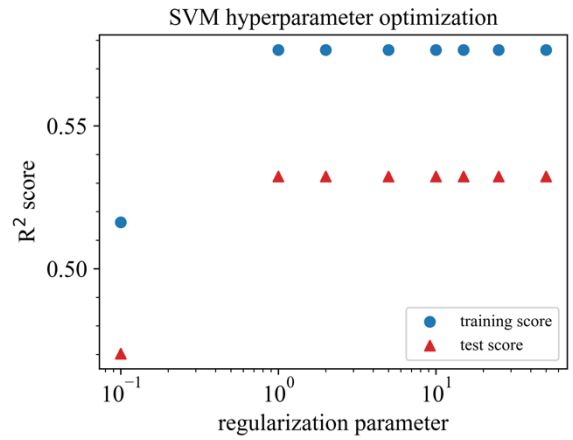

Figure S191: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, ϵ = 1.0, descriptor set A, and density as target property.

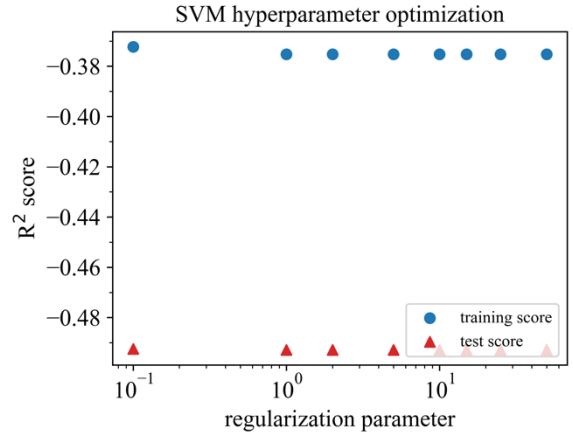

Figure S192: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, ϵ = 2.0, descriptor set A, and density as target property.

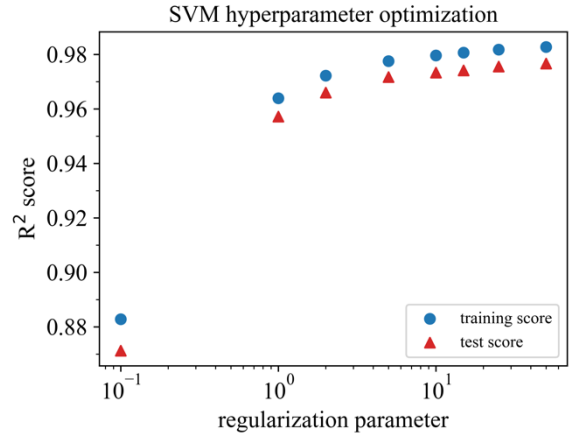

Figure S193: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, ϵ = 0.01, descriptor set C, and density as target property.

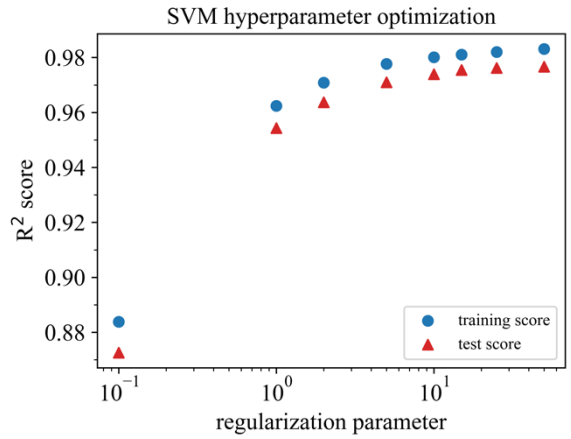

Figure S194: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.1$, descriptor set C, and density as target property.

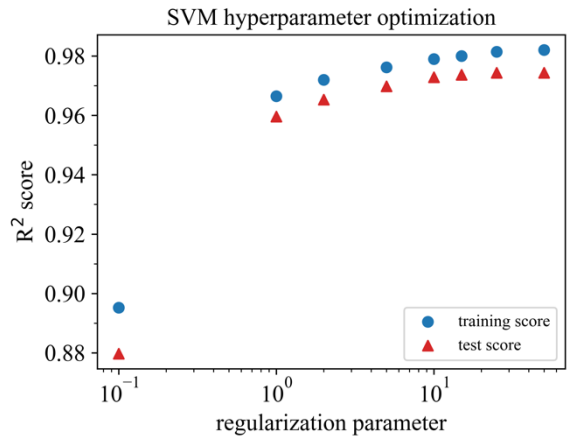

Figure S195: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.2$, descriptor set C, and density as target property.

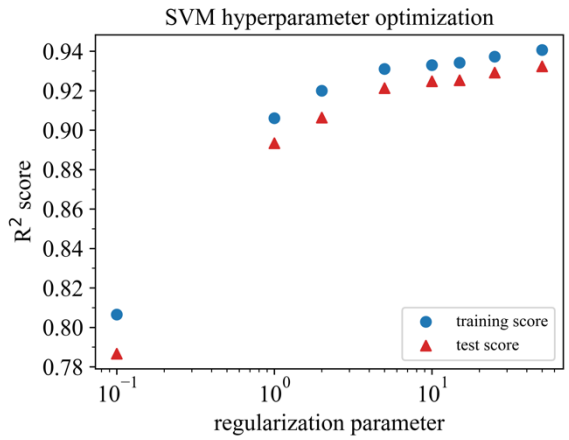

Figure S196: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.5$, descriptor set C, and density as target property.

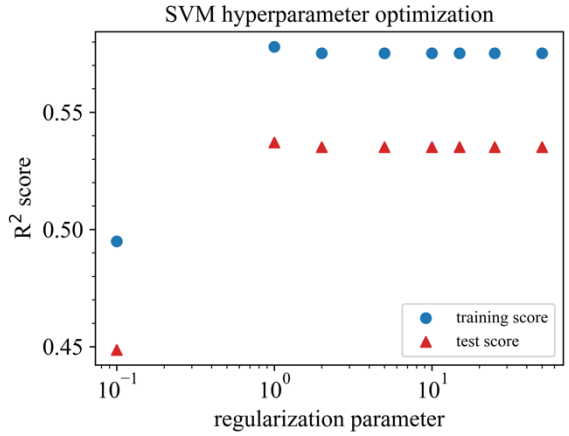

Figure S197: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 1.0$, descriptor set C, and density as target property.

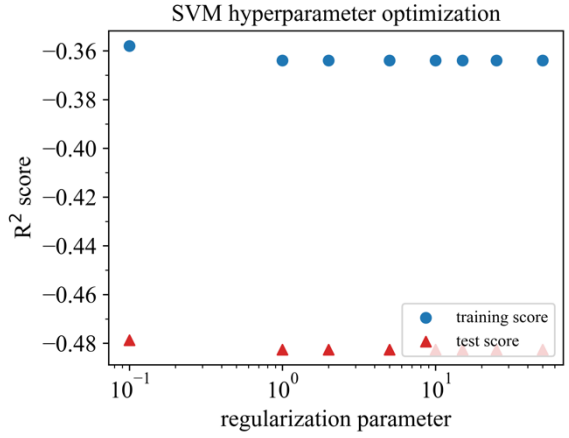

Figure S198: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 2.0$, descriptor set C, and density as target property.

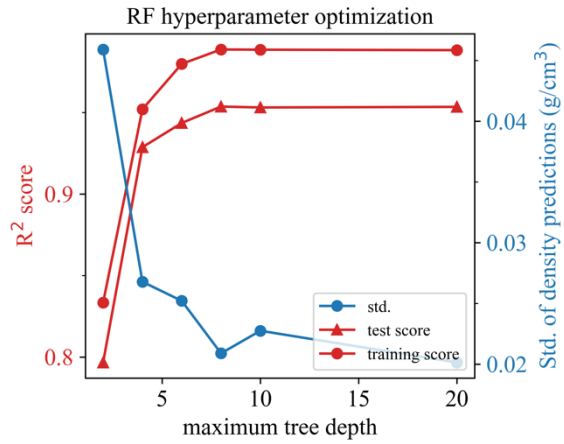

Figure S199: Hyperparameter optimization for RF trained with the V14 dataset, 4 estimators, descriptor set AC, and density as target property.

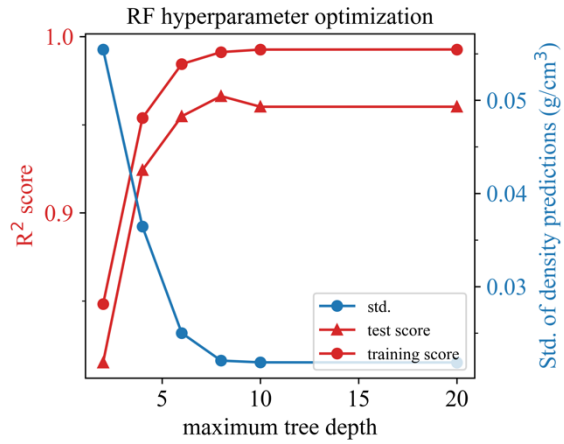

Figure S200: Hyperparameter optimization for RF trained with the V14 dataset, 8 estimators, descriptor set AC, and density as target property.

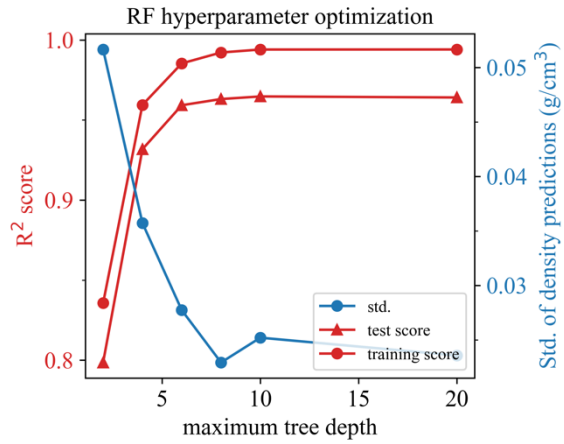

Figure S201: Hyperparameter optimization for RF trained with the V14 dataset, 16 estimators, descriptor set AC, and density as target property.

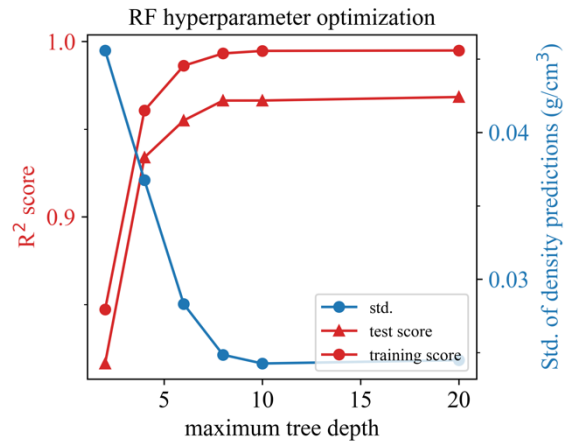

Figure S202: Hyperparameter optimization for RF trained with the V14 dataset, 32 estimators, descriptor set AC, and density as target property.

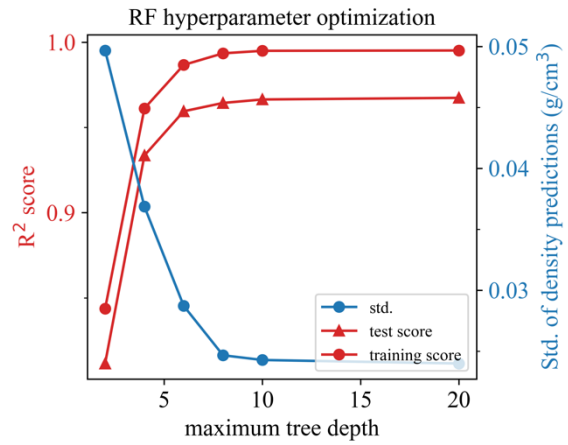

Figure S203: Hyperparameter optimization for RF trained with the V14 dataset, 64 estimators, descriptor set AC, and density as target property.

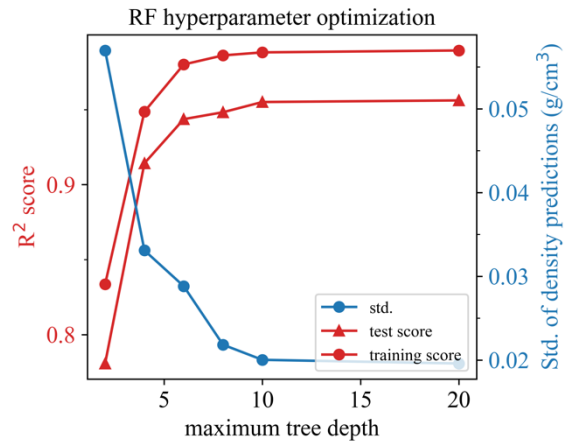

Figure S204: Hyperparameter optimization for RF trained with the V14 dataset, 4 estimators, descriptor set A, and density as target property.

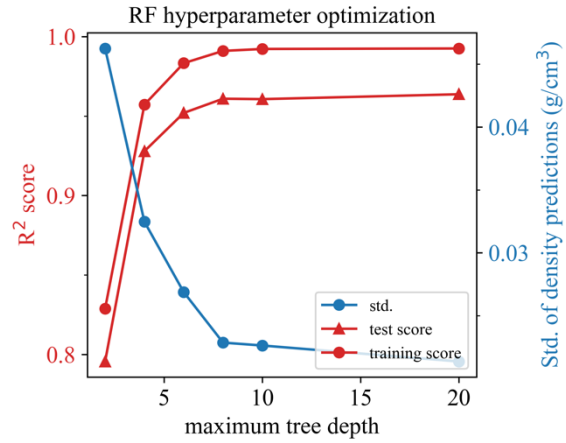

Figure S205: Hyperparameter optimization for RF trained with the V14 dataset, 8 estimators, descriptor set A, and density as target property.

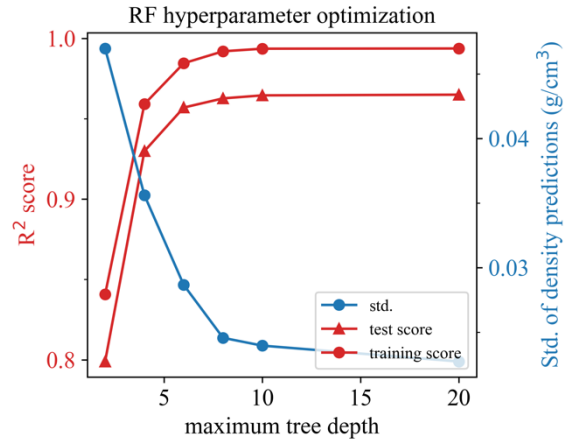

Figure S206: Hyperparameter optimization for RF trained with the V14 dataset, 16 estimators, descriptor set A, and density as target property.

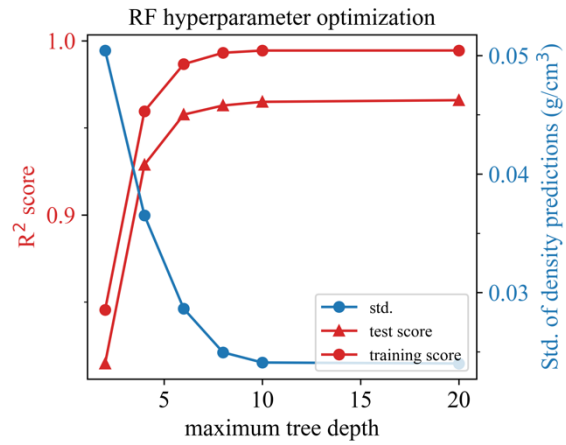

Figure S207: Hyperparameter optimization for RF trained with the V14 dataset, 32 estimators, descriptor set A, and density as target property.

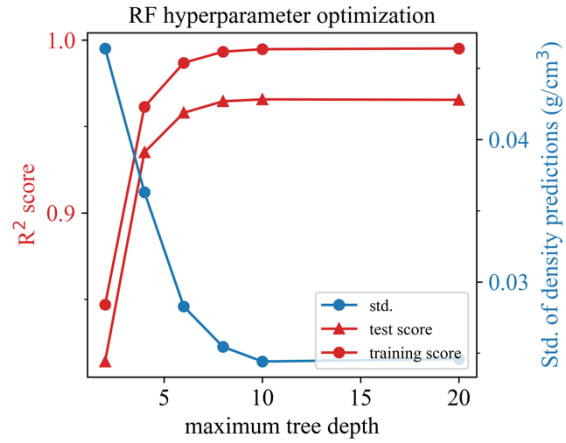

Figure S208: Hyperparameter optimization for RF trained with the V14 dataset, 64 estimators, descriptor set A, and density as target property.

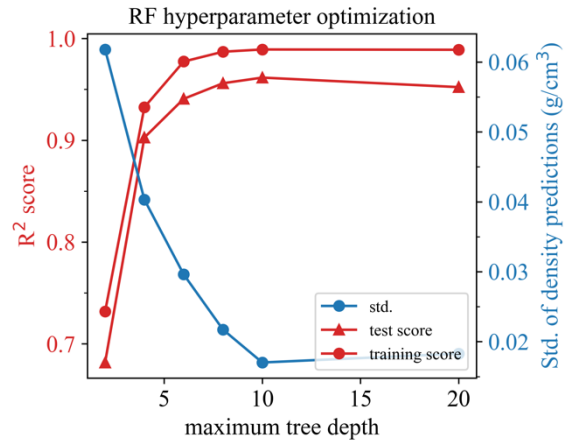

Figure S209: Hyperparameter optimization for RF trained with the V14 dataset, 4 estimators, descriptor set C, and density as target property.

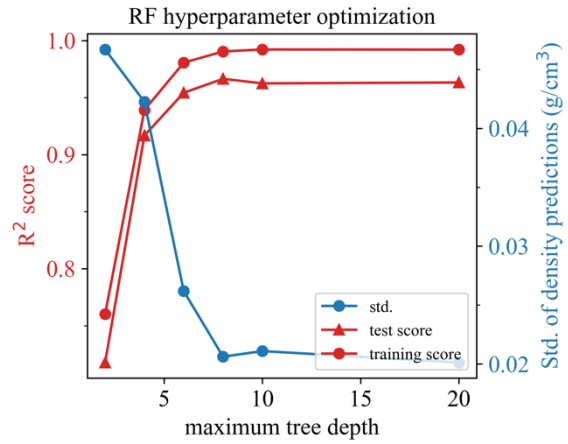

Figure S210: Hyperparameter optimization for RF trained with the V14 dataset, 8 estimators, descriptor set C, and density as target property.

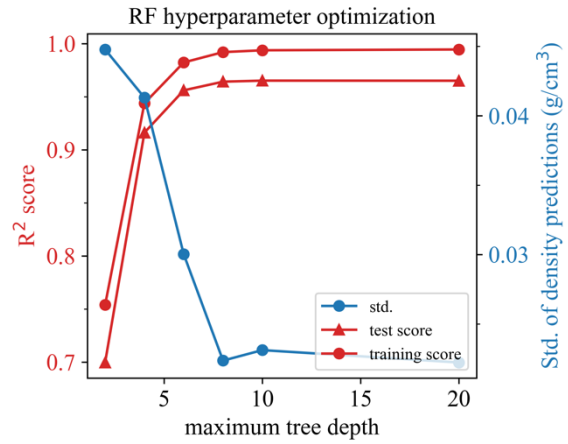

Figure S211: Hyperparameter optimization for RF trained with the V14 dataset, 16 estimators, descriptor set C, and density as target property.

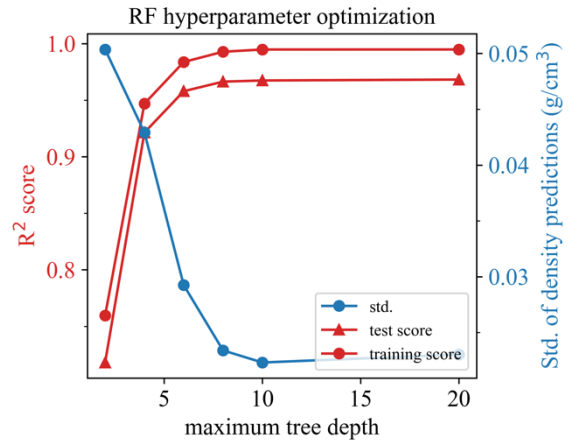

Figure S212: Hyperparameter optimization for RF trained with the V14 dataset, 32 estimators, descriptor set C, and density as target property.

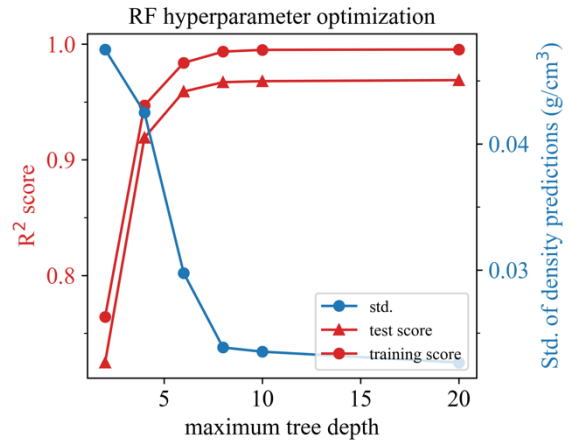

Figure S213: Hyperparameter optimization for RF trained with the V14 dataset, 64 estimators, descriptor set C, and density as target property.

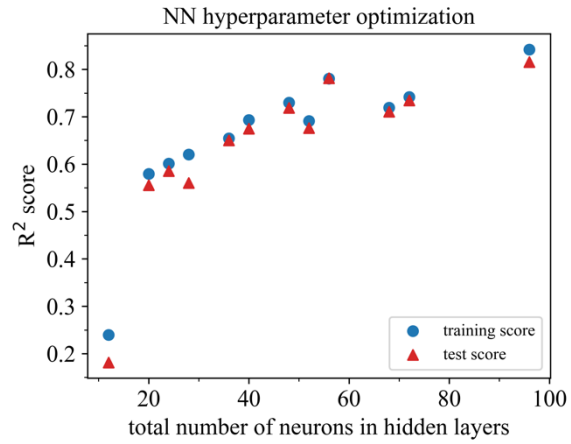

Figure S214: Hyperparameter optimization for NN trained with the V14 dataset, descriptor set AC, and density as target property.

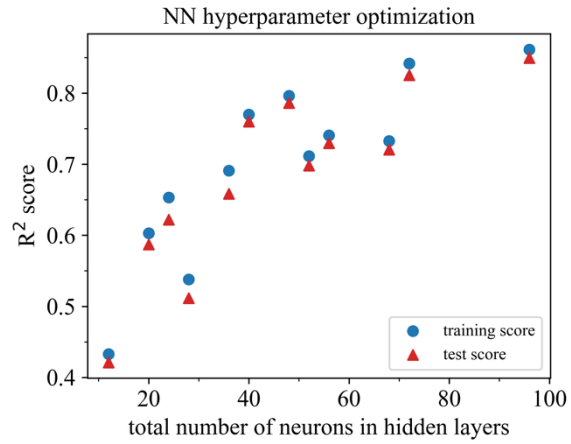

Figure S215: Hyperparameter optimization for NN trained with the V14 dataset, descriptor set A, and density as target property.

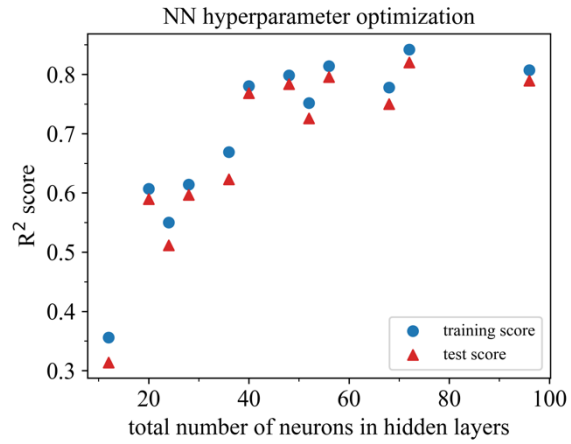

Figure S216: Hyperparameter optimization for NN trained with the V14 dataset, descriptor set C, and density as target property.

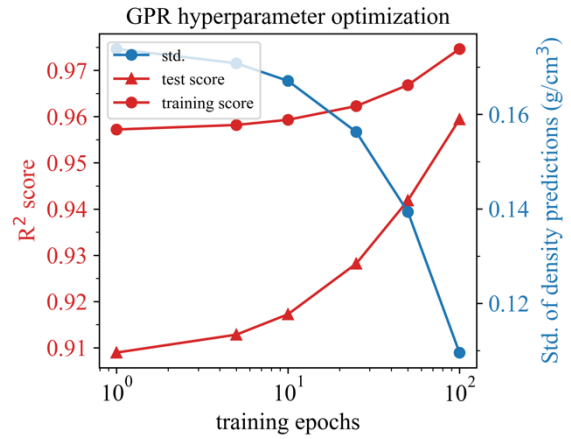

Figure S217: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.01, descriptor set AC, and density as target property.

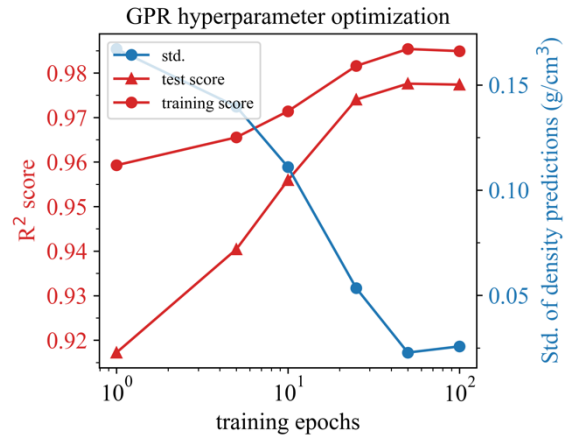

Figure S218: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.1, descriptor set AC, and density as target property.

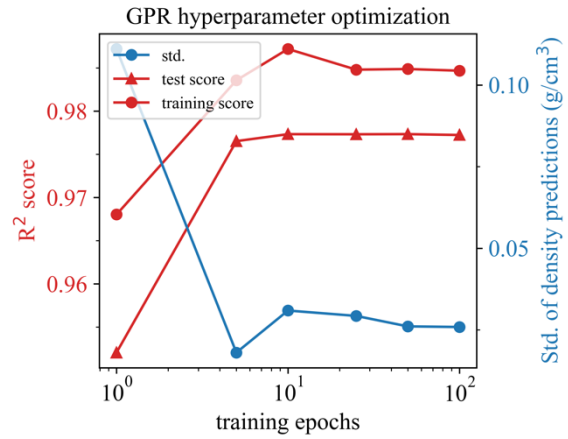

Figure S219: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 1.0, descriptor set AC, and density as target property.

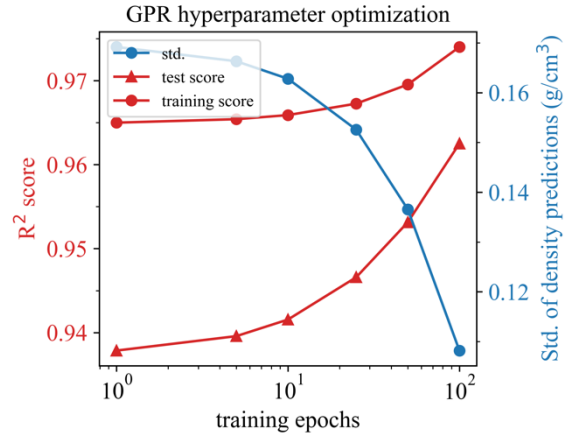

Figure S220: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.01, descriptor set A, and density as target property.

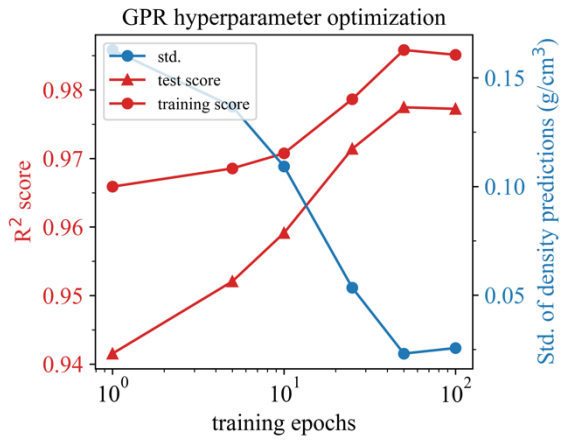

Figure S221: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.1, descriptor set A, and density as target property.

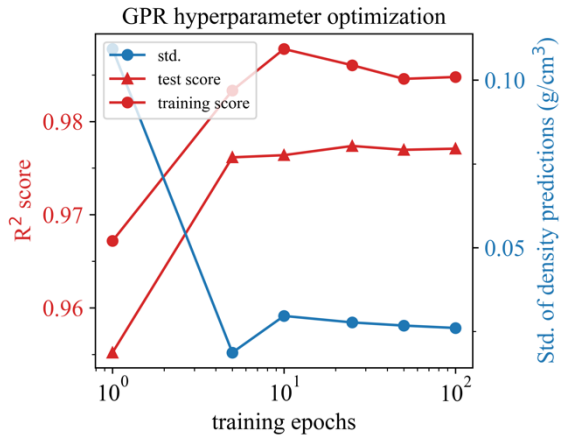

Figure S222: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 1.0, descriptor set A, and density as target property.

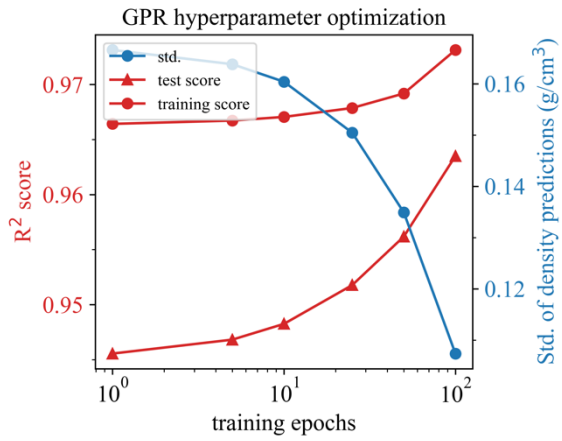

Figure S223: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.01, descriptor set C, and density as target property.

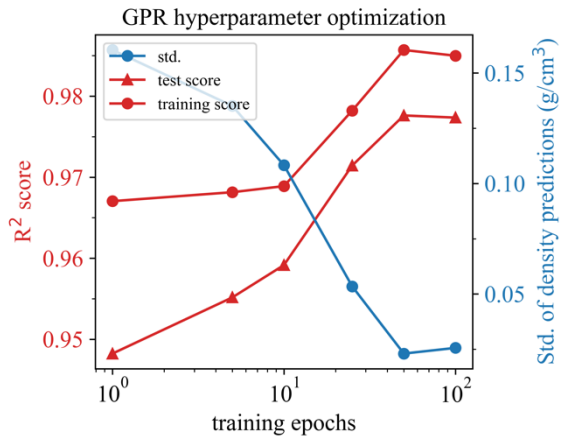

Figure S224: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.1, descriptor set C, and density as target property.

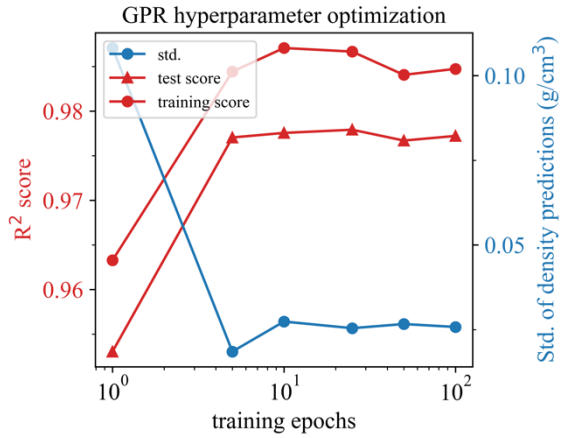

Figure S225: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 1.0, descriptor set C, and density as target property.

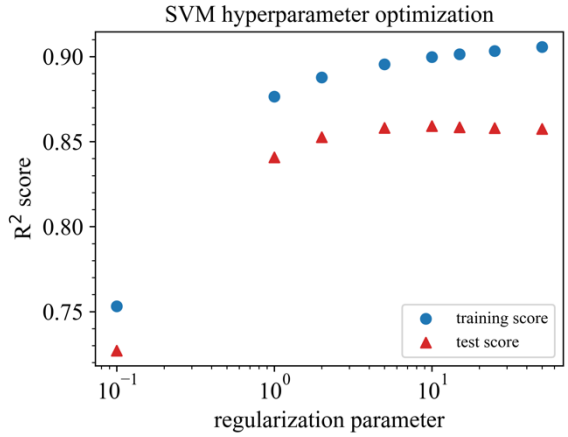

Figure S226: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.01$, descriptor set AC, and bulk modulus as target property.

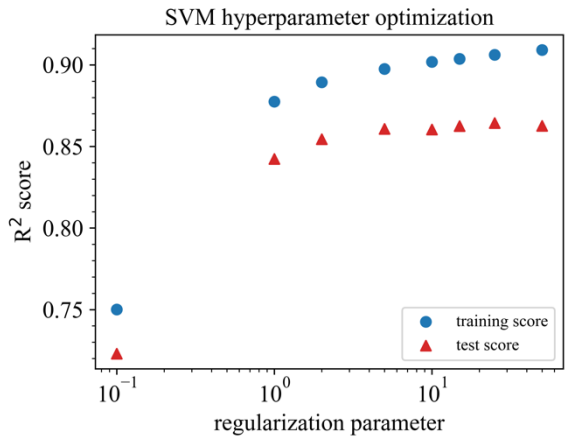

Figure S227: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.1$, descriptor set AC, and bulk modulus as target property.

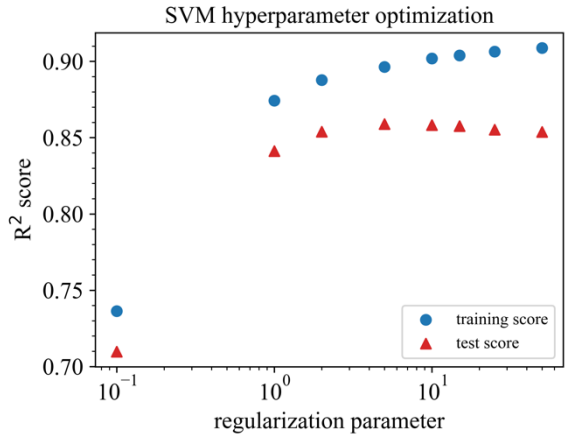

Figure S228: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.2$, descriptor set AC, and bulk modulus as target property.

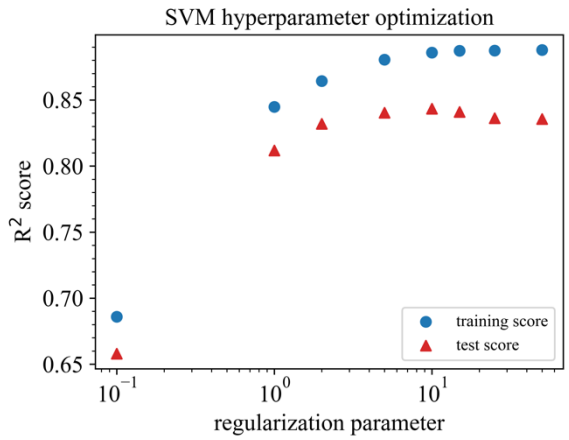

Figure S229: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.5$, descriptor set AC, and bulk modulus as target property.

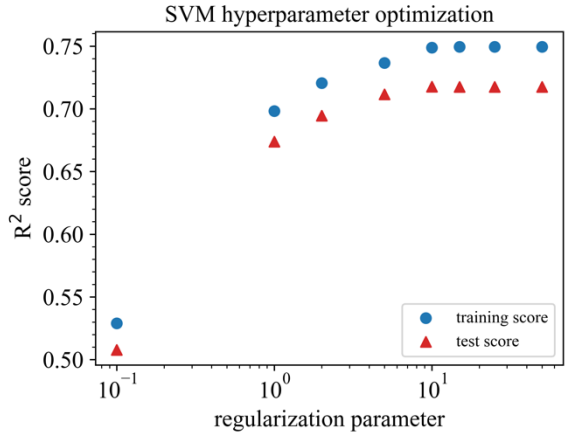

Figure S230: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 1.0$, descriptor set AC, and bulk modulus as target property.

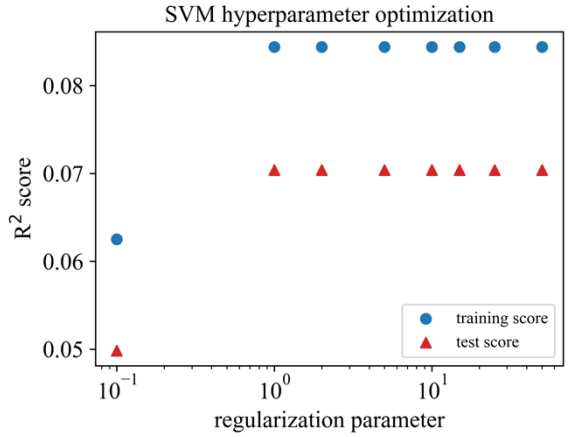

Figure S231: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 2.0$, descriptor set AC, and bulk modulus as target property.

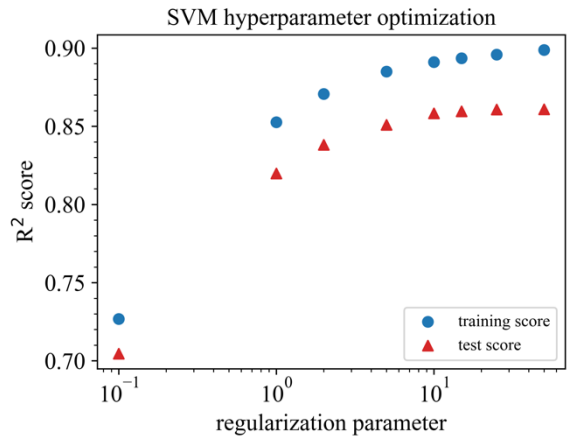

Figure S232: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.01$, descriptor set A, and bulk modulus as target property.

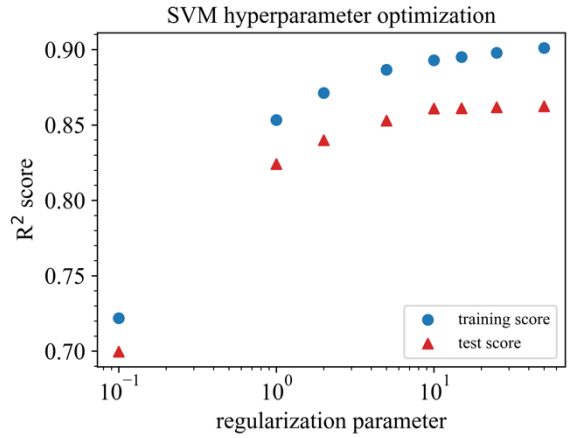

Figure S233: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.1$, descriptor set A, and bulk modulus as target property.

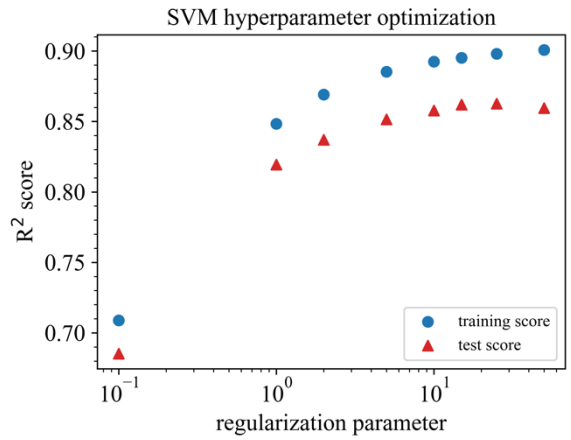

Figure S234: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.2$, descriptor set A, and bulk modulus as target property.

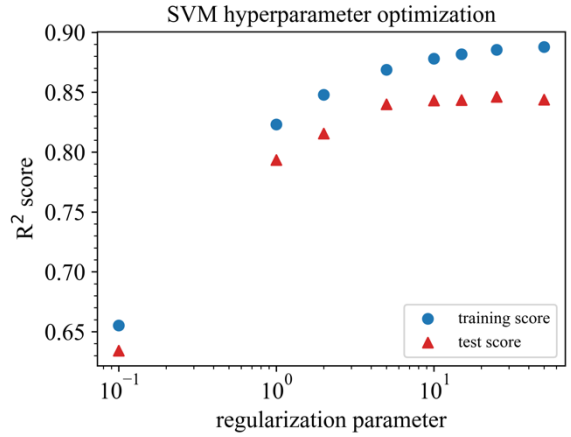

Figure S235: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.5$, descriptor set A, and bulk modulus as target property.

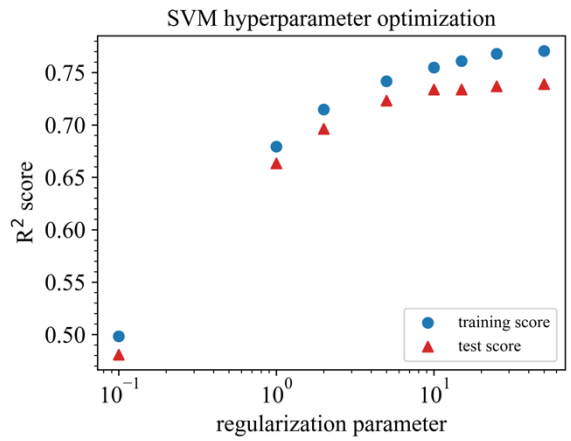

Figure S236: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 1.0$, descriptor set A, and bulk modulus as target property.

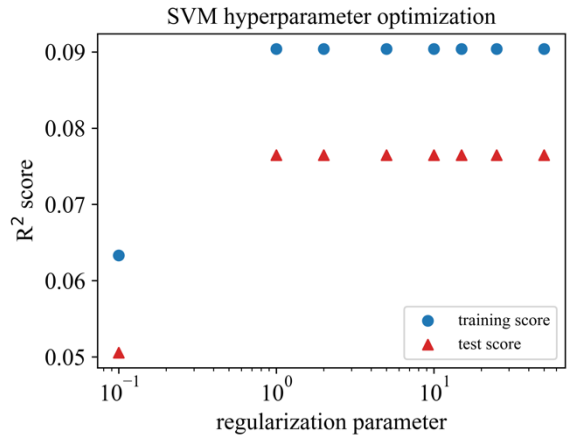

Figure S237: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 2.0$, descriptor set A, and bulk modulus as target property.

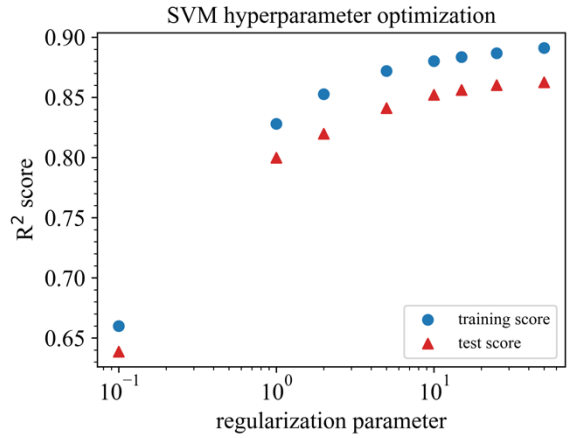

Figure S238: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.01$, descriptor set C, and bulk modulus as target property.

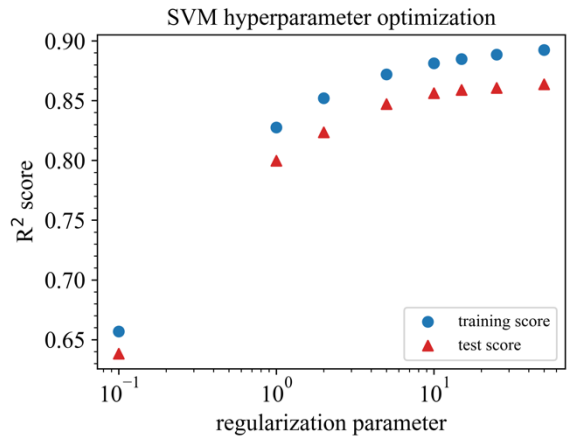

Figure S239: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.1$, descriptor set C, and bulk modulus as target property.

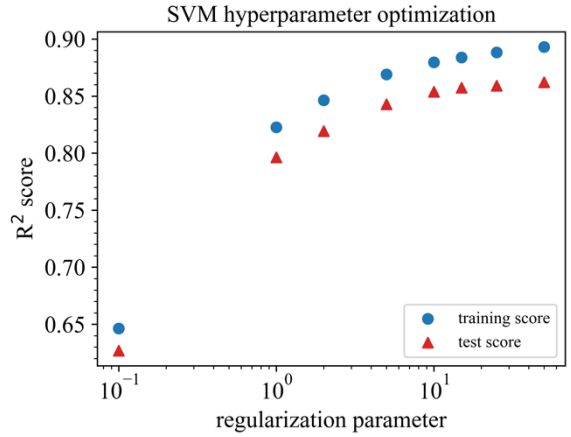

Figure S240: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.2$, descriptor set C, and bulk modulus as target property.

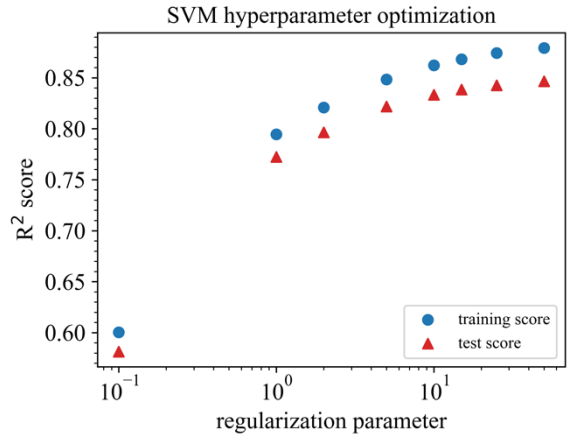

Figure S241: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.5$, descriptor set C, and bulk modulus as target property.

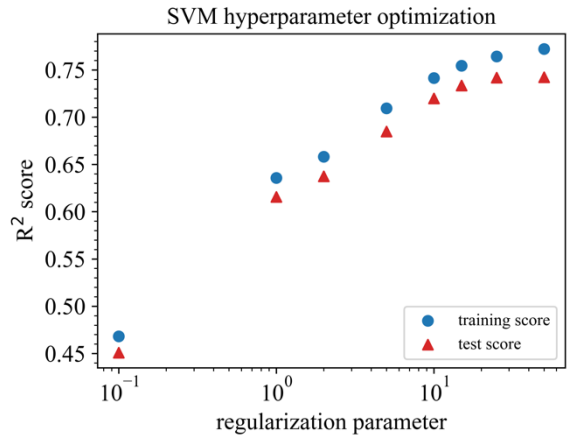

Figure S242: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, ϵ = 1.0, descriptor set C, and bulk modulus as target property.

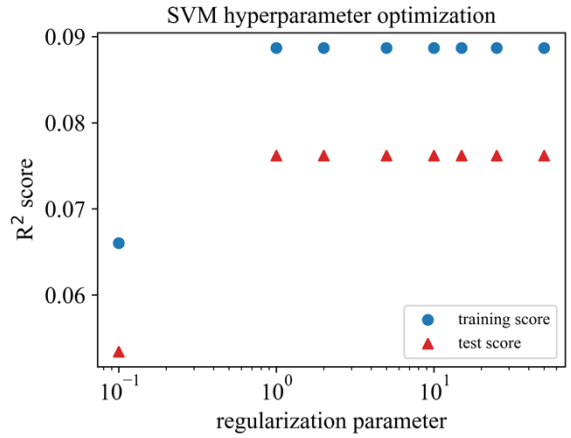

Figure S243: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, ϵ = 2.0, descriptor set C, and bulk modulus as target property.

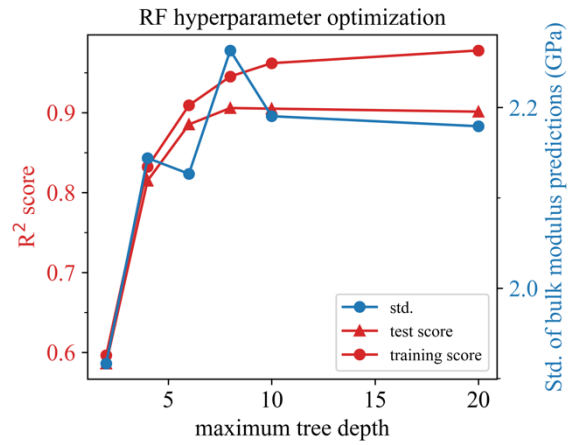

Figure S244: Hyperparameter optimization for RF trained with the V14 dataset, 4 estimators, descriptor set AC, and bulk modulus as target property.

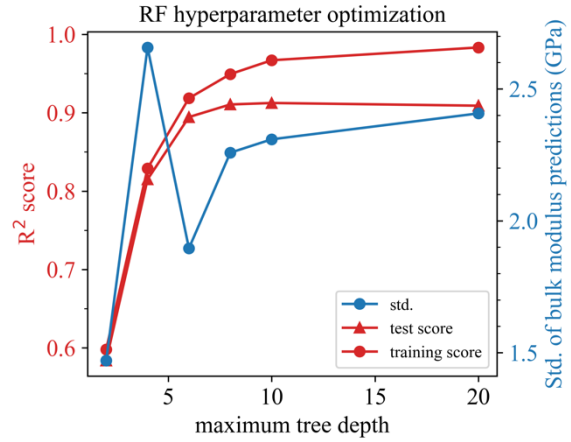

Figure S245: Hyperparameter optimization for RF trained with the V14 dataset, 8 estimators, descriptor set AC, and bulk modulus as target property.

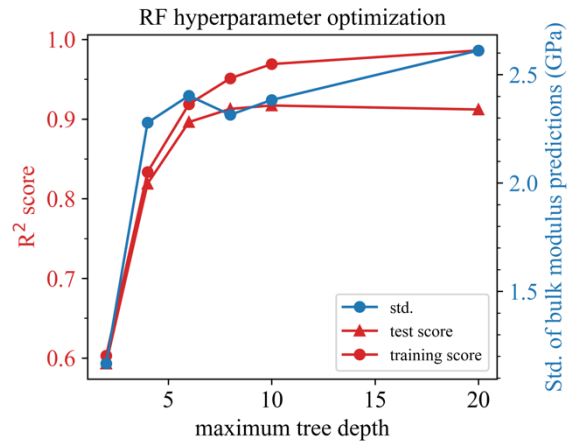

Figure S246: Hyperparameter optimization for RF trained with the V14 dataset, 16 estimators, descriptor set AC, and bulk modulus as target property.

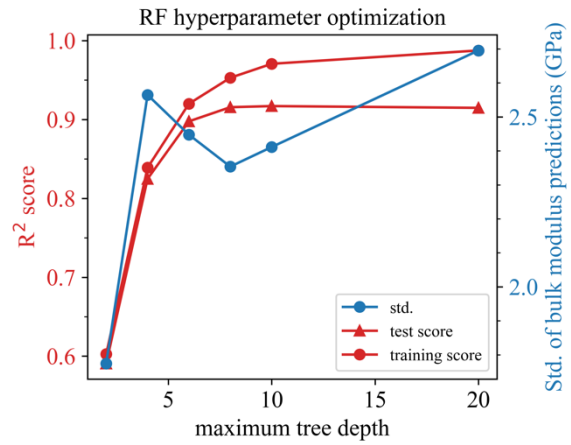

Figure S247: Hyperparameter optimization for RF trained with the V14 dataset, 32 estimators, descriptor set AC, and bulk modulus as target property.

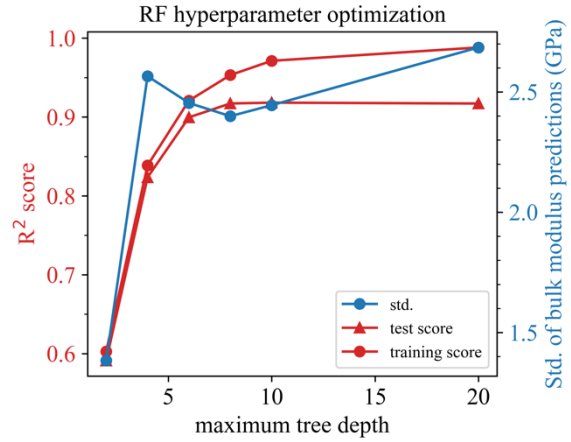

Figure S248: Hyperparameter optimization for RF trained with the V14 dataset, 64 estimators, descriptor set AC, and bulk modulus as target property.

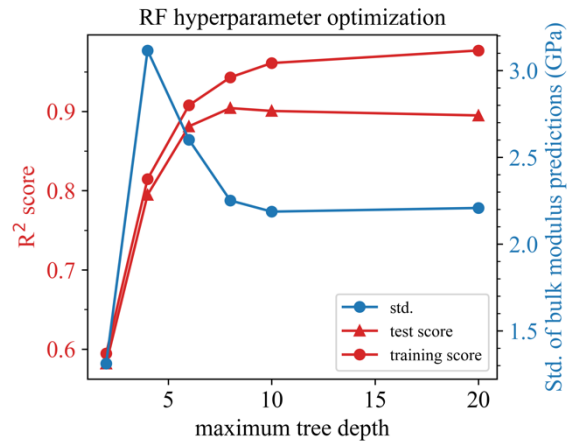

Figure S249: Hyperparameter optimization for RF trained with the V14 dataset, 4 estimators, descriptor set A, and bulk modulus as target property.

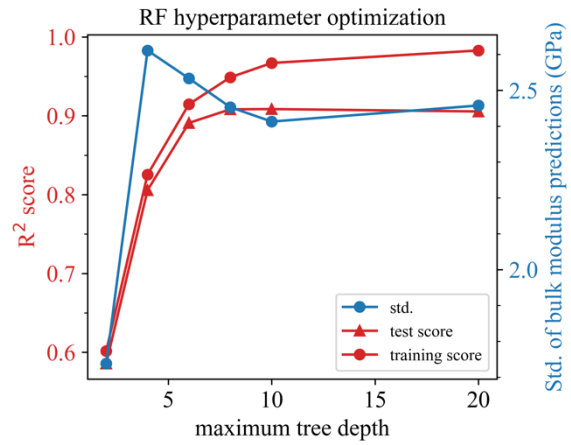

Figure S250: Hyperparameter optimization for RF trained with the V14 dataset, 8 estimators, descriptor set A, and bulk modulus as target property.

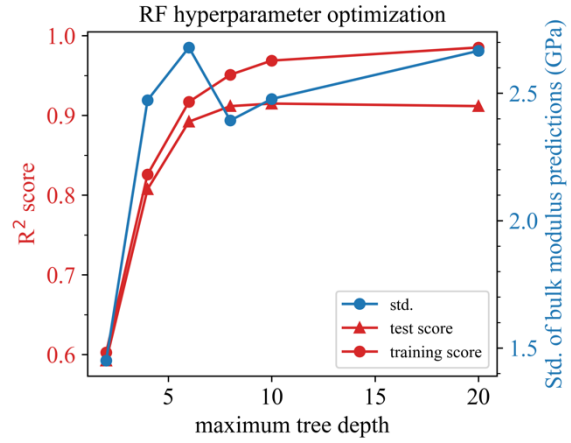

Figure S251: Hyperparameter optimization for RF trained with the V14 dataset, 16 estimators, descriptor set A, and bulk modulus as target property.

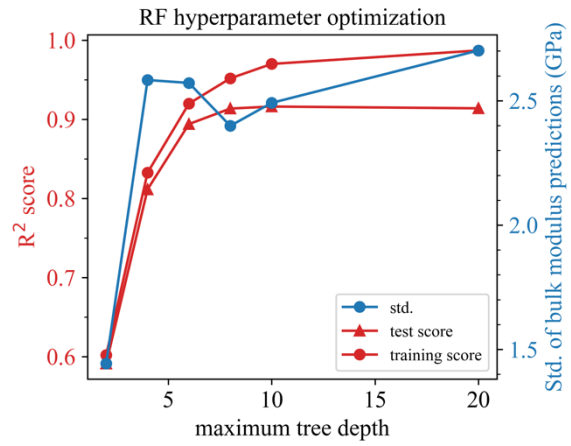

Figure S252: Hyperparameter optimization for RF trained with the V14 dataset, 32 estimators, descriptor set A, and bulk modulus as target property.

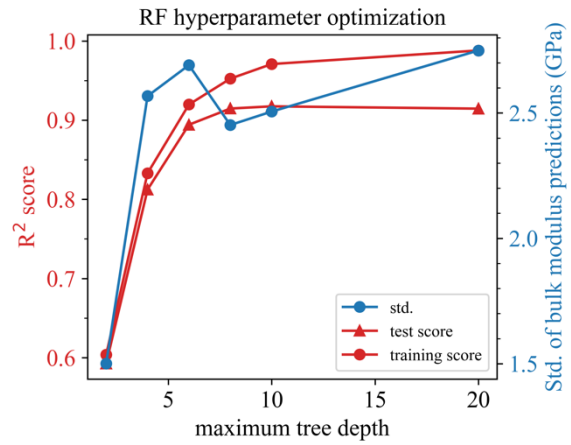

Figure S253: Hyperparameter optimization for RF trained with the V14 dataset, 64 estimators, descriptor set A, and bulk modulus as target property.

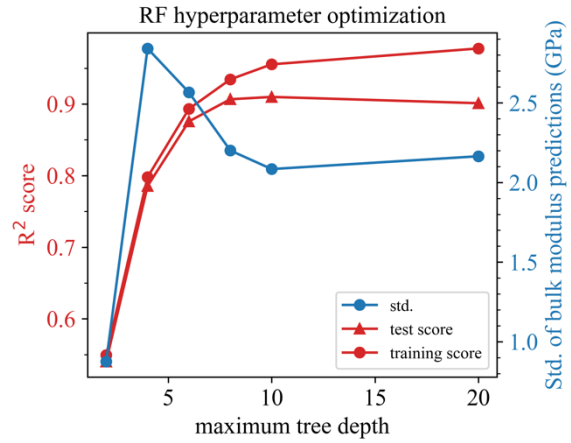

Figure S254: Hyperparameter optimization for RF trained with the V14 dataset, 4 estimators, descriptor set C, and bulk modulus as target property.

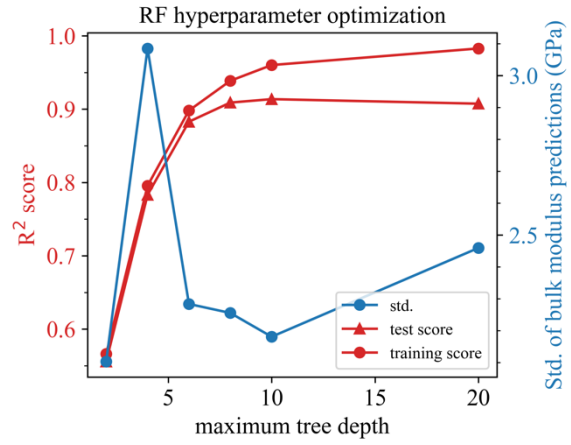

Figure S255: Hyperparameter optimization for RF trained with the V14 dataset, 8 estimators, descriptor set C, and bulk modulus as target property.

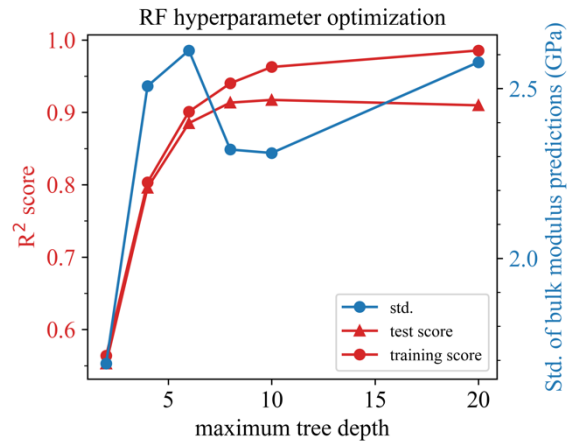

Figure S256: Hyperparameter optimization for RF trained with the V14 dataset, 16 estimators, descriptor set C, and bulk modulus as target property.

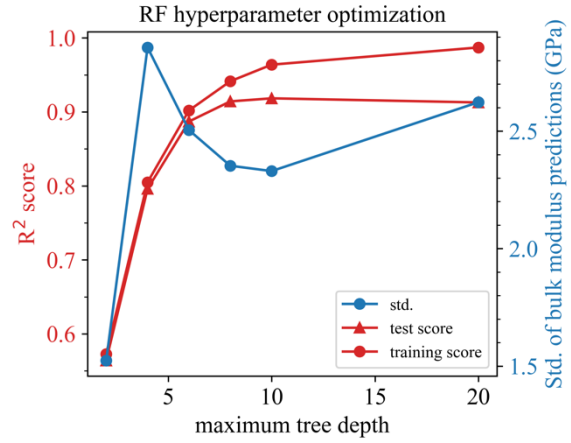

Figure S257: Hyperparameter optimization for RF trained with the V14 dataset, 32 estimators, descriptor set C, and bulk modulus as target property.

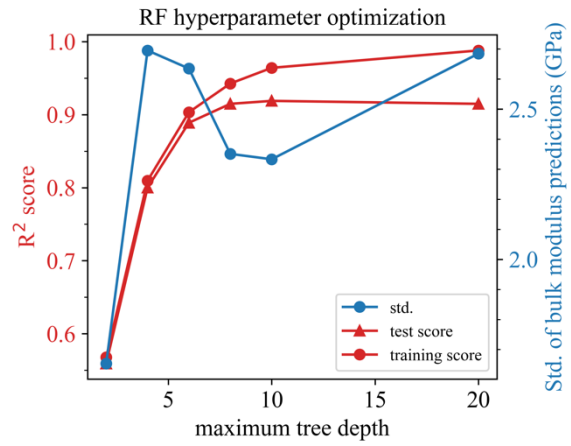

Figure S258: Hyperparameter optimization for RF trained with the V14 dataset, 64 estimators, descriptor set C, and bulk modulus as target property.

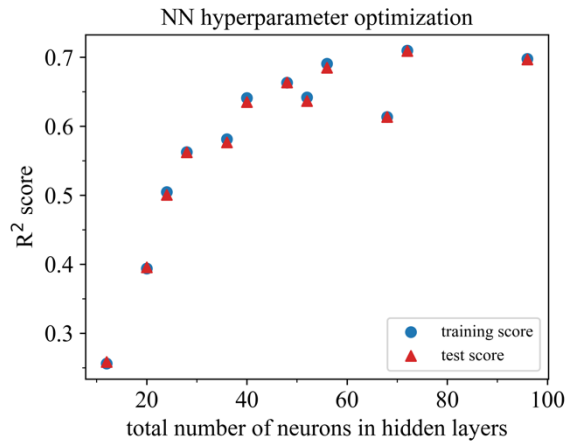

Figure S259: Hyperparameter optimization for NN trained with the V14 dataset, descriptor set AC, and bulk modulus as target property.

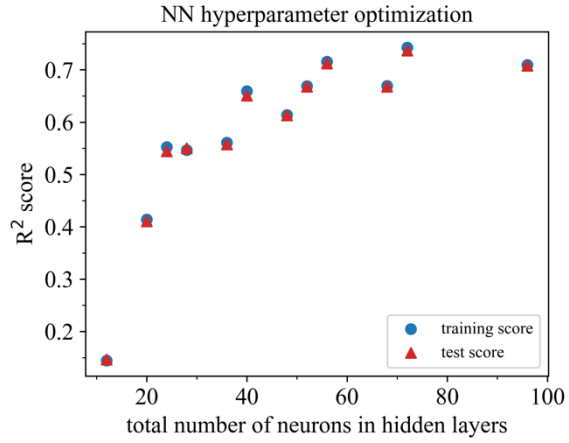

Figure S260: Hyperparameter optimization for NN trained with the V14 dataset, descriptor set A, and bulk modulus as target property.

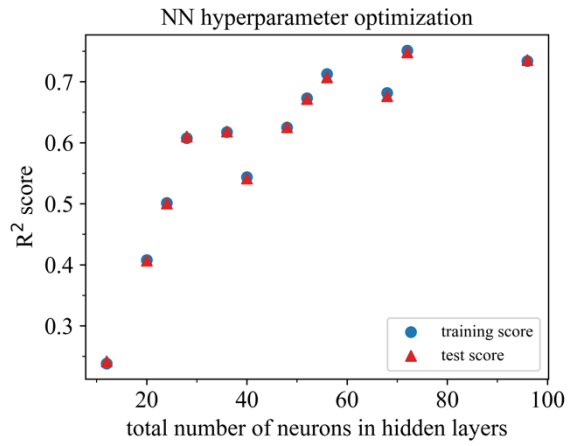

Figure S261: Hyperparameter optimization for NN trained with the V14 dataset, descriptor set C, and bulk modulus as target property.

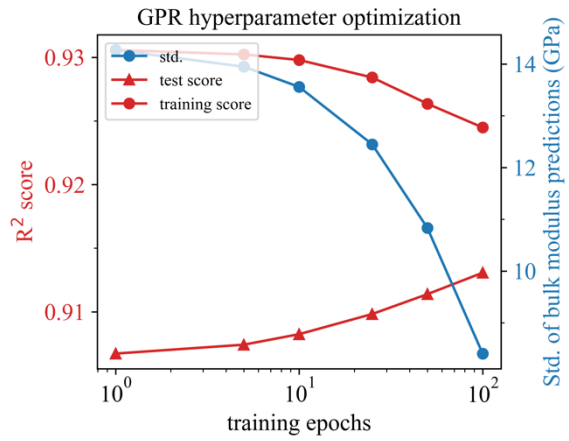

Figure S262: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.01, descriptor set AC, and bulk modulus as target property.

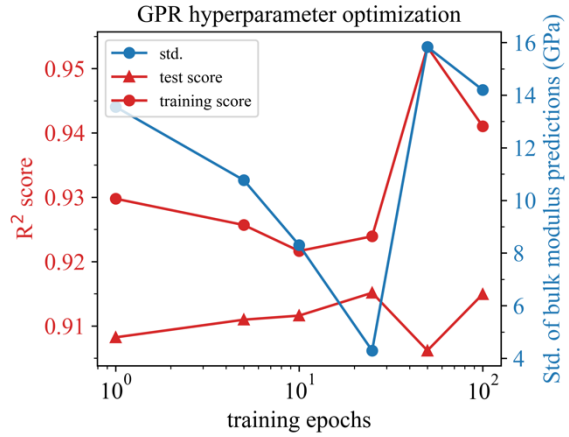

Figure S263: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.1, descriptor set AC, and bulk modulus as target property.

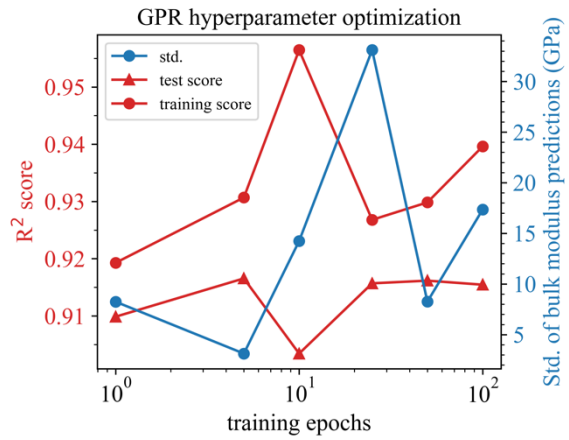

Figure S264: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 1.0, descriptor set AC, and bulk modulus as target property.

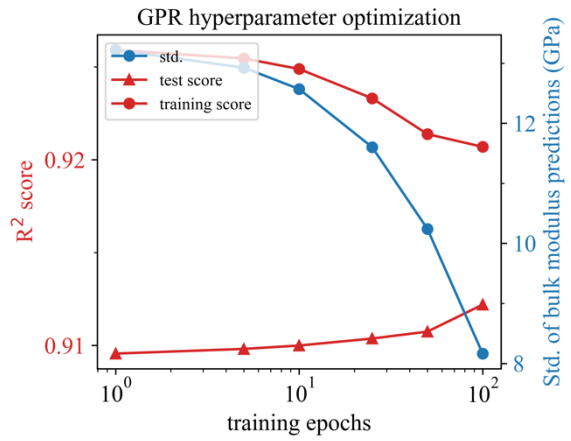

Figure S265: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.01, descriptor set A, and bulk modulus as target property.

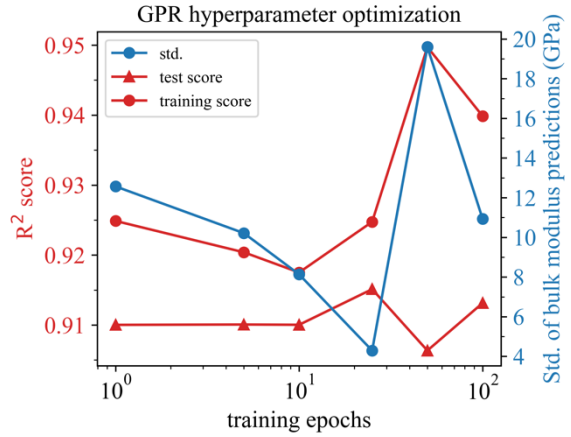

Figure S266: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.1, descriptor set A, and bulk modulus as target property.

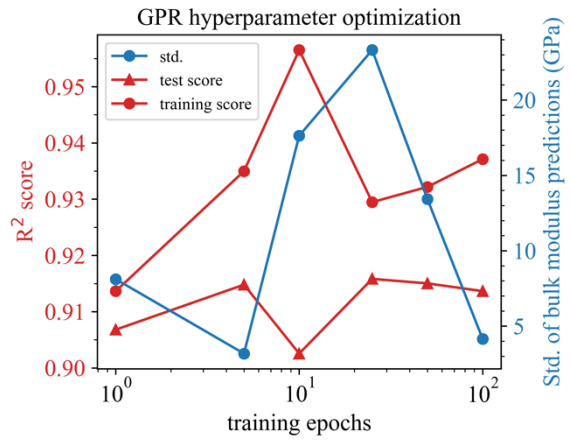

Figure S267: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 1.0, descriptor set A, and bulk modulus as target property.

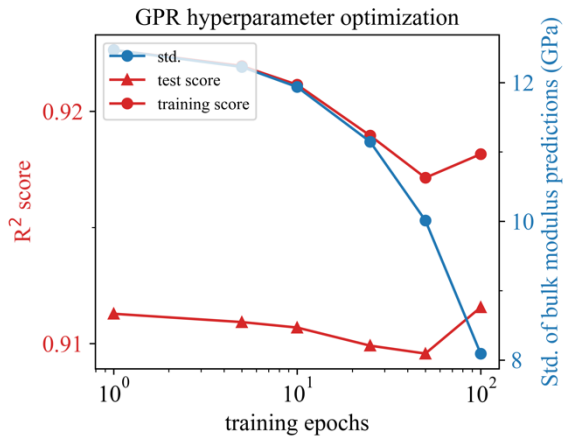

Figure S268: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.01, descriptor set C, and bulk modulus as target property.

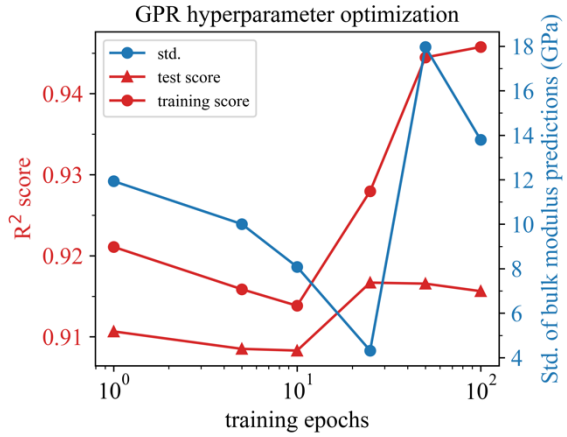

Figure S269: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.1, descriptor set C, and bulk modulus as target property.

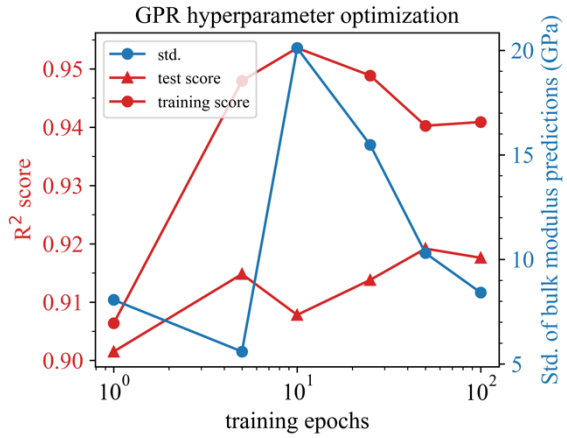

Figure S270: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 1.0, descriptor set C, and bulk modulus as target property.

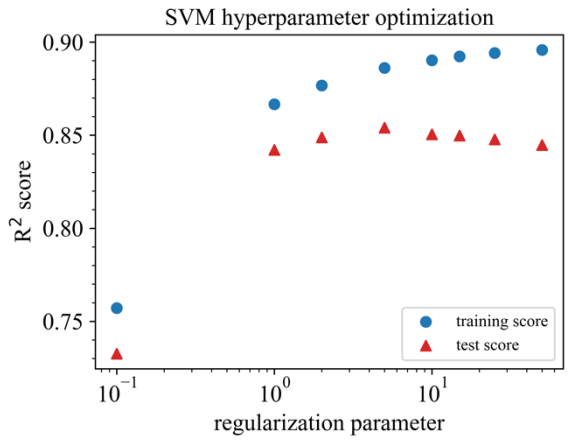

Figure S271: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.01$, descriptor set AC, and shear modulus as target property.

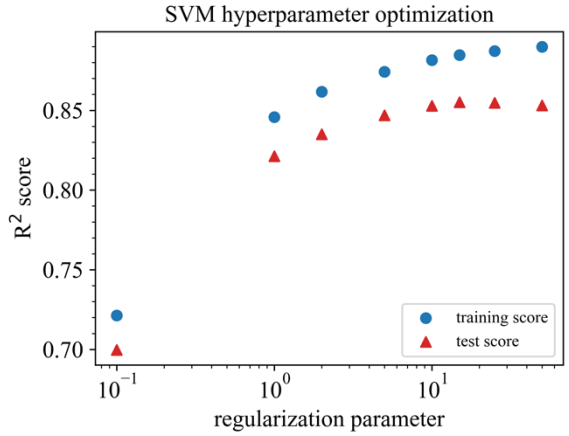

Figure S272: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.1$, descriptor set AC, and shear modulus as target property.

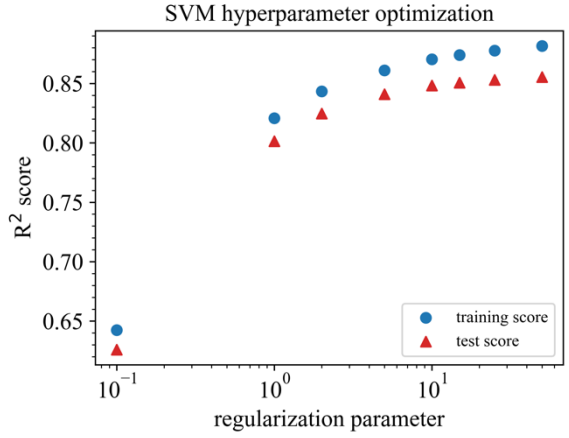

Figure S273: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.2$, descriptor set AC, and shear modulus as target property.

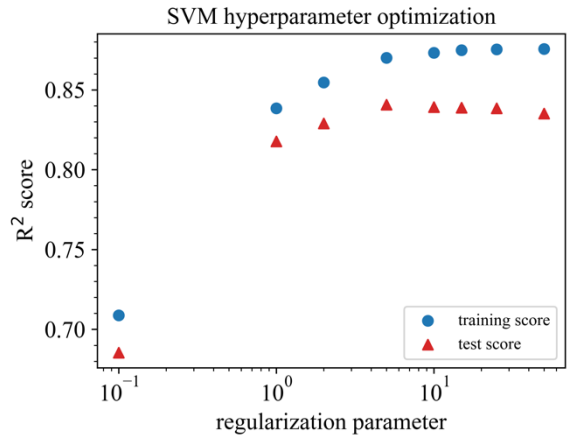

Figure S274: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.5$, descriptor set AC, and shear modulus as target property.

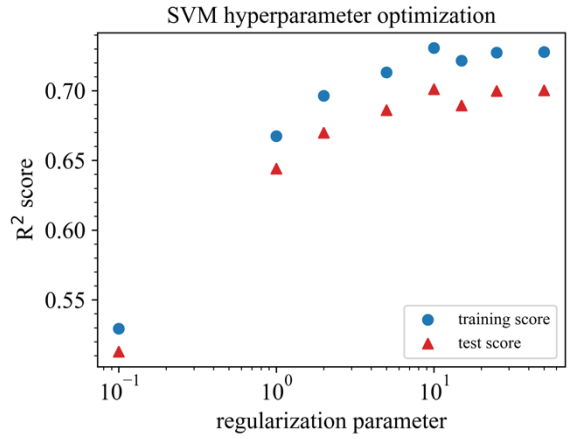

Figure S275: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 1.0$, descriptor set AC, and shear modulus as target property.

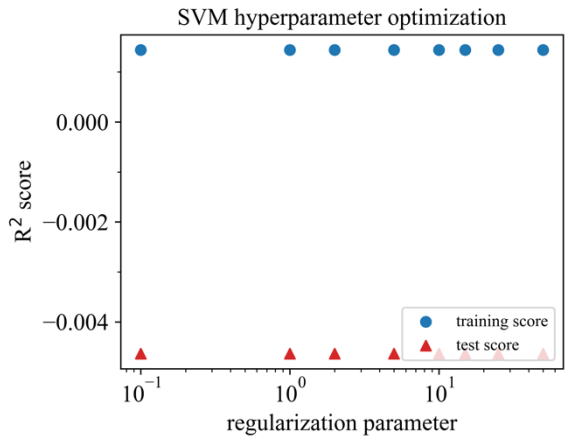

Figure S276: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 2.0$, descriptor set AC, and shear modulus as target property.

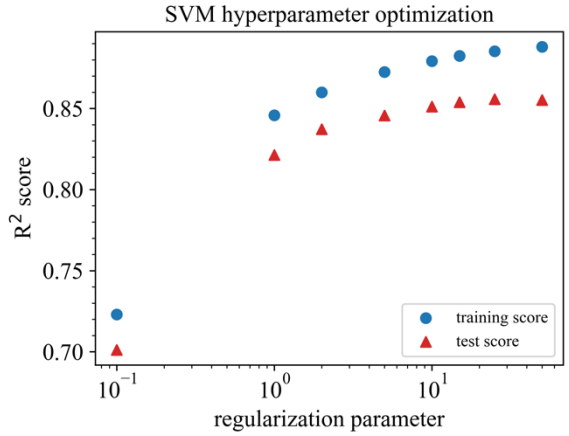

Figure S277: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.01$, descriptor set A, and shear modulus as target property.

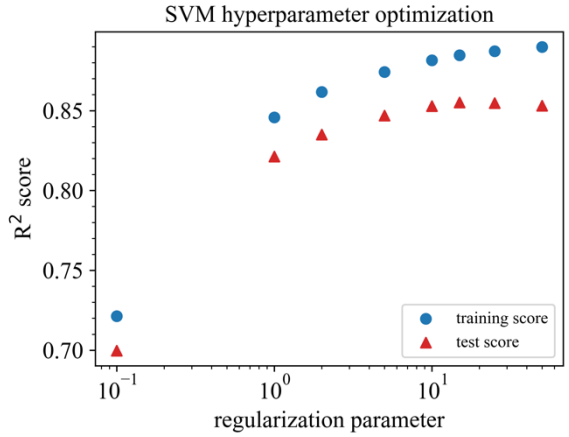

Figure S278: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.1$, descriptor set A, and shear modulus as target property.

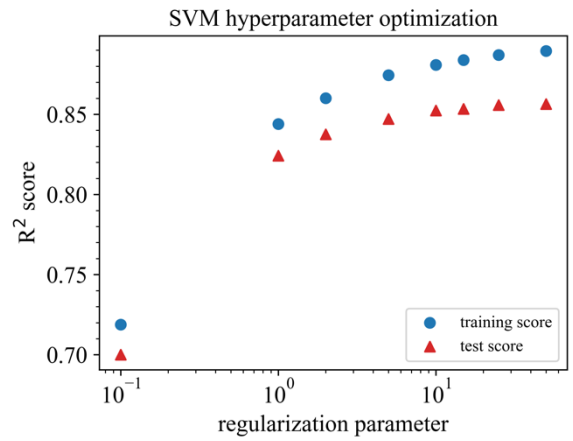

Figure S279: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.2$, descriptor set A, and shear modulus as target property.

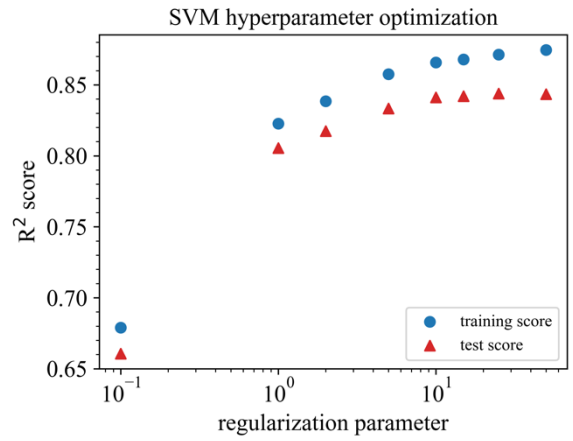

Figure S280: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.5$, descriptor set A, and shear modulus as target property.

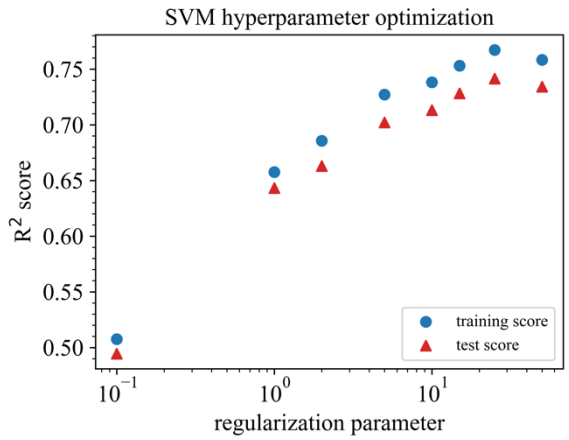

Figure S281: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, ϵ = 1.0, descriptor set A, and shear modulus as target property.

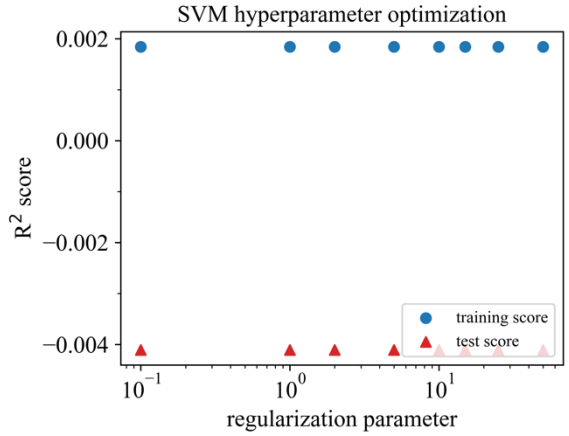

Figure S282: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, ϵ = 2.0, descriptor set A, and shear modulus as target property.

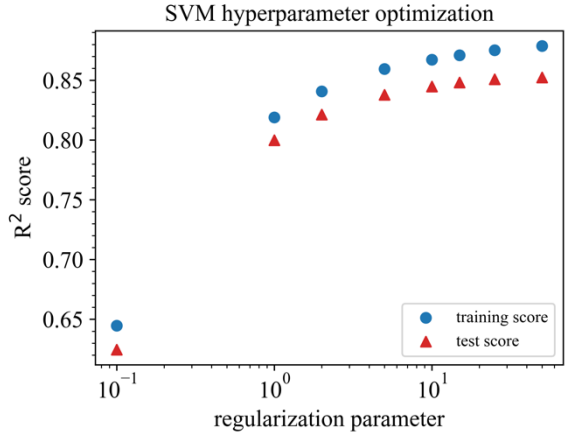

Figure S283: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, ϵ = 0.01, descriptor set C, and shear modulus as target property.

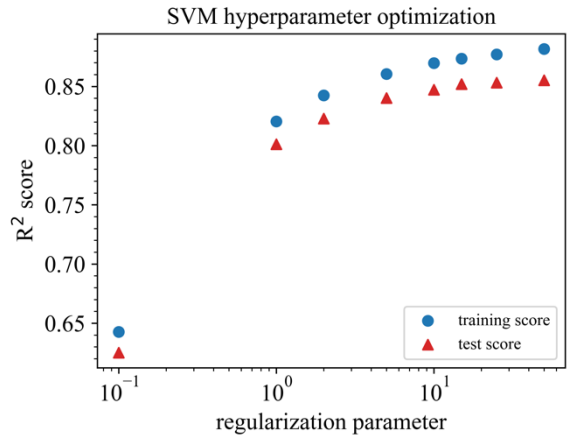

Figure S284: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.1$, descriptor set C, and shear modulus as target property.

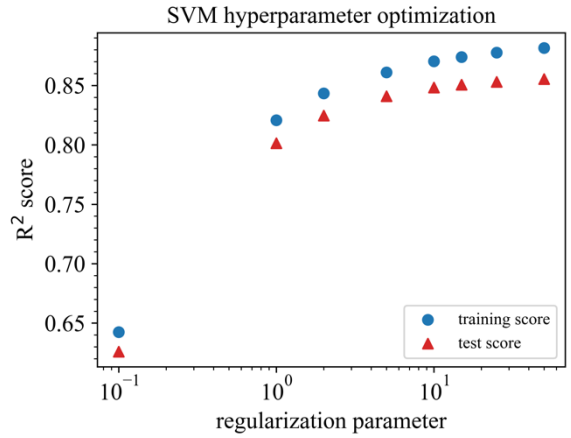

Figure S285: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.2$, descriptor set C, and shear modulus as target property.

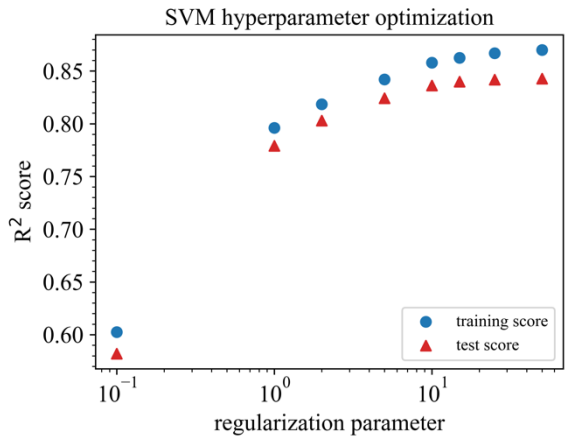

Figure S286: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.5$, descriptor set C, and shear modulus as target property.

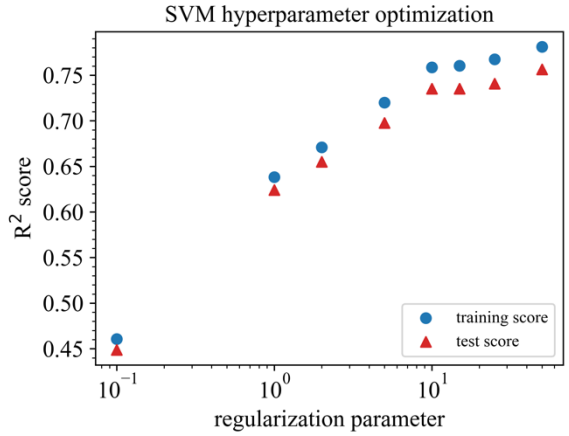

Figure S287: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 1.0$, descriptor set C, and shear modulus as target property.

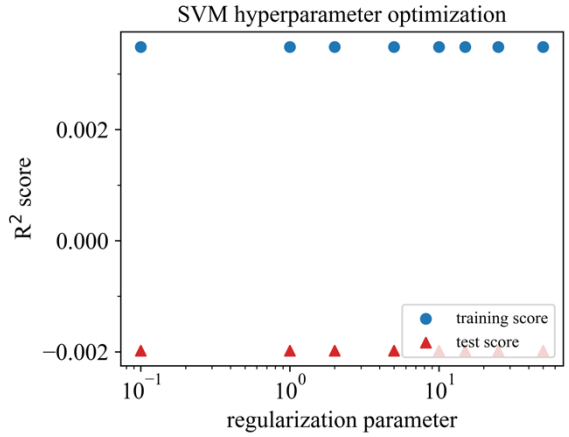

Figure S288: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 2.0$, descriptor set C, and shear modulus as target property.

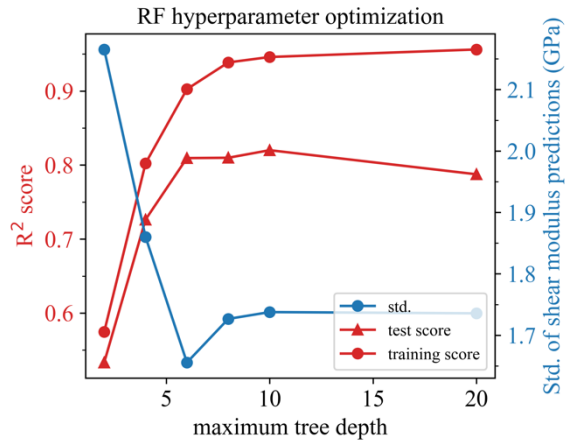

Figure S289: Hyperparameter optimization for RF trained with the V14 dataset, 4 estimators, descriptor set AC, and shear modulus as target property.

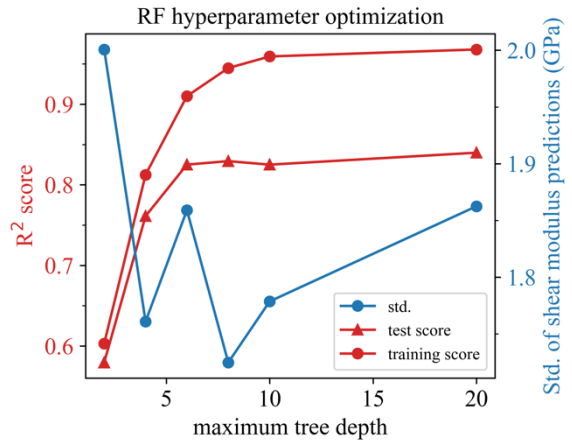

Figure S290: Hyperparameter optimization for RF trained with the V14 dataset, 8 estimators, descriptor set AC, and shear modulus as target property.

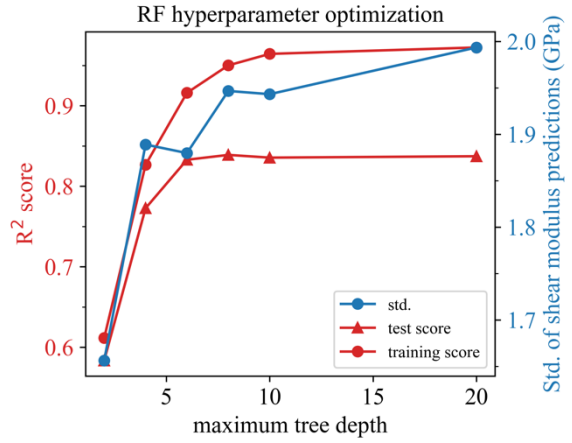

Figure S291: Hyperparameter optimization for RF trained with the V14 dataset, 16 estimators, descriptor set AC, and shear modulus as target property.

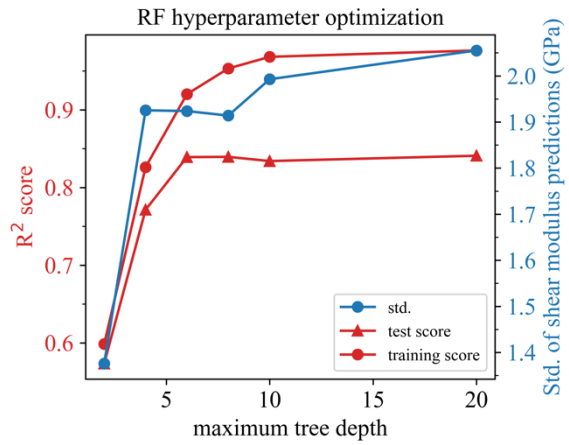

Figure S292: Hyperparameter optimization for RF trained with the V14 dataset, 32 estimators, descriptor set AC, and shear modulus as target property.

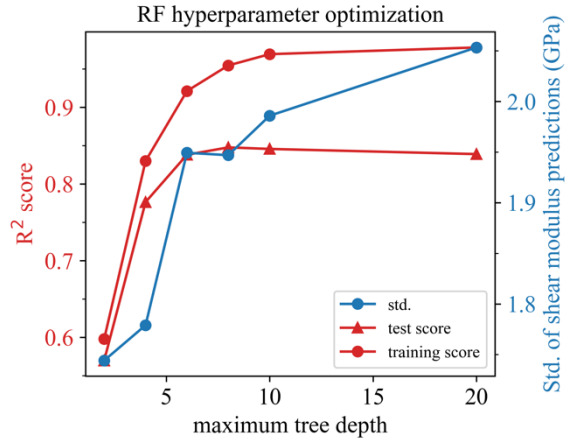

Figure S293: Hyperparameter optimization for RF trained with the V14 dataset, 64 estimators, descriptor set AC, and shear modulus as target property.

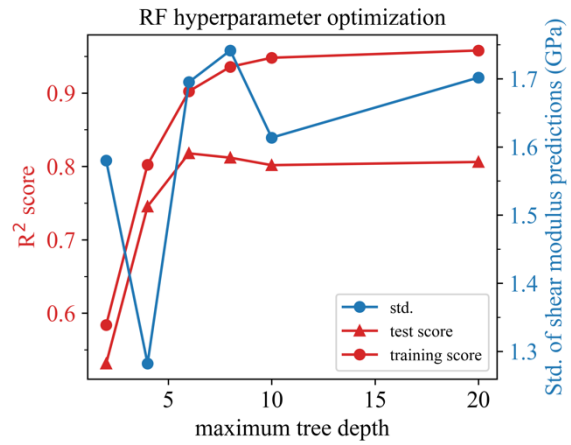

Figure S294: Hyperparameter optimization for RF trained with the V14 dataset, 4 estimators, descriptor set A, and shear modulus as target property.

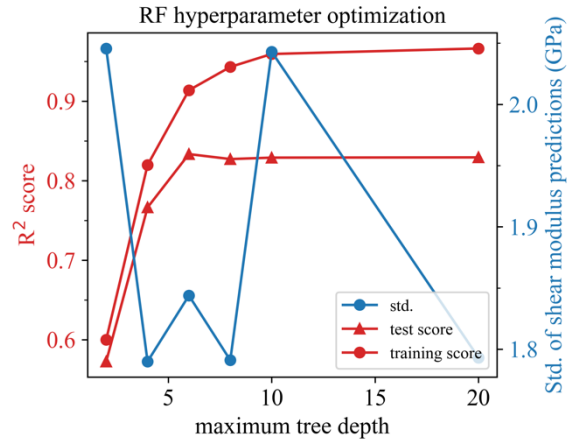

Figure S295: Hyperparameter optimization for RF trained with the V14 dataset, 8 estimators, descriptor set A, and shear modulus as target property.

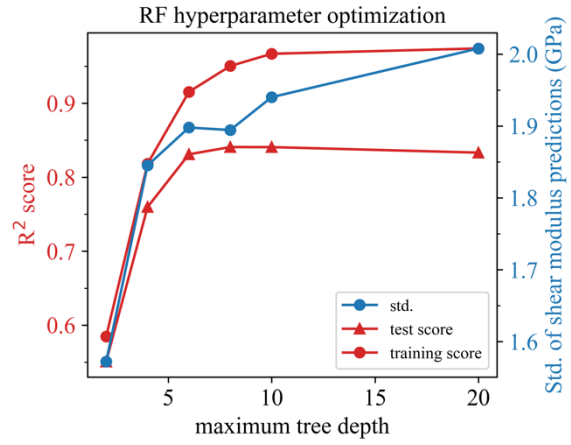

Figure S296: Hyperparameter optimization for RF trained with the V14 dataset, 16 estimators, descriptor set A, and shear modulus as target property.

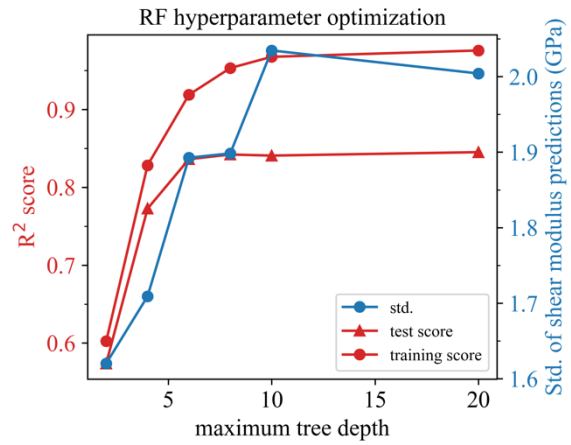

Figure S297: Hyperparameter optimization for RF trained with the V14 dataset, 32 estimators, descriptor set A, and shear modulus as target property.

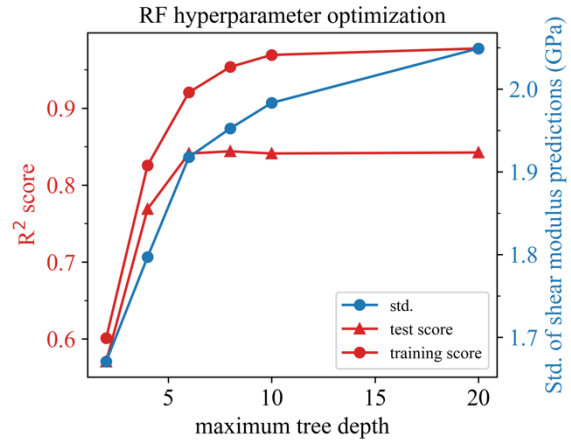

Figure S298: Hyperparameter optimization for RF trained with the V14 dataset, 64 estimators, descriptor set A, and shear modulus as target property.

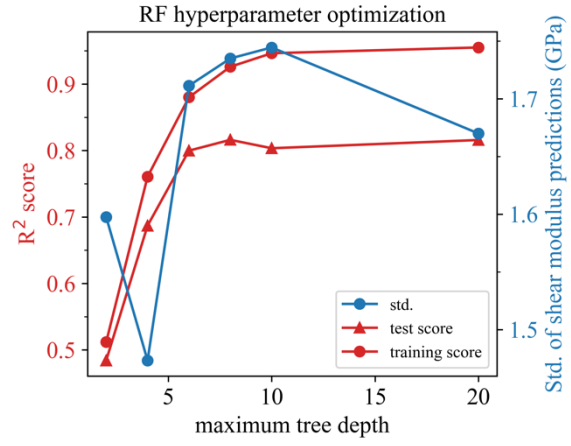

Figure S299: Hyperparameter optimization for RF trained with the V14 dataset, 4 estimators, descriptor set C, and shear modulus as target property.

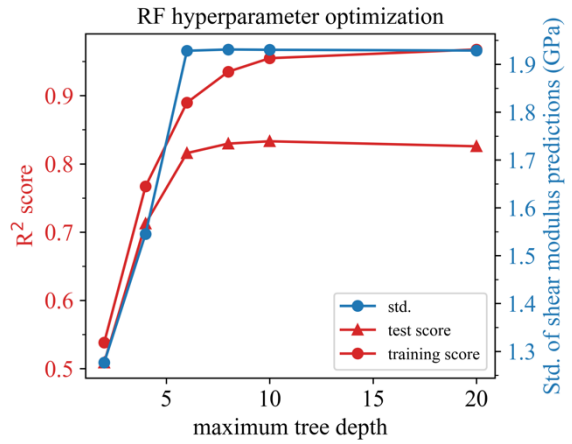

Figure S300: Hyperparameter optimization for RF trained with the V14 dataset, 8 estimators, descriptor set C, and shear modulus as target property.

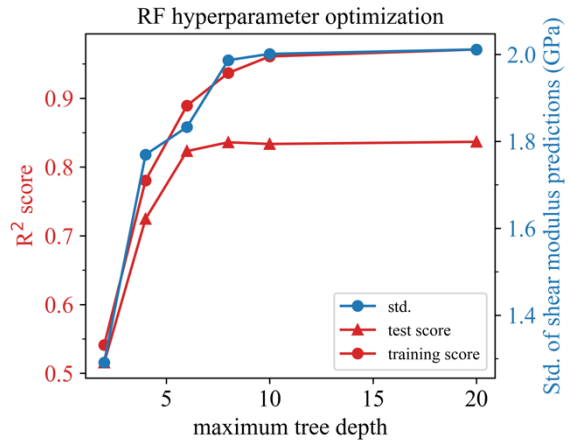

Figure S301: Hyperparameter optimization for RF trained with the V14 dataset, 16 estimators, descriptor set C, and shear modulus as target property.

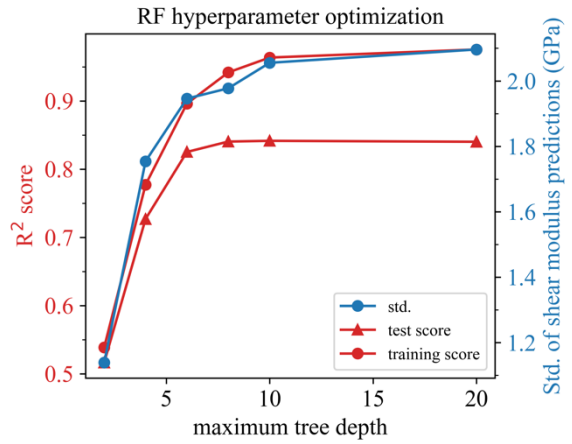

Figure S302: Hyperparameter optimization for RF trained with the V14 dataset, 32 estimators, descriptor set C, and shear modulus as target property.

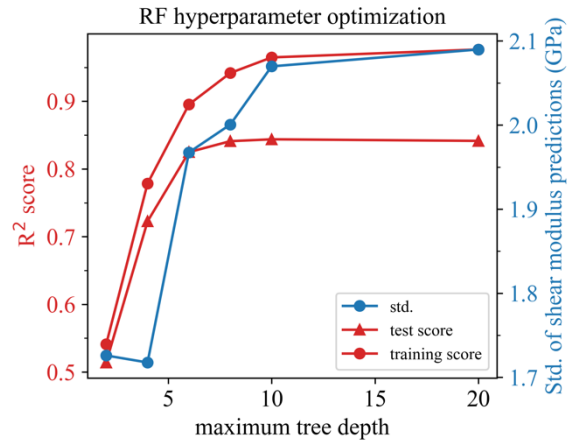

Figure S303: Hyperparameter optimization for RF trained with the V14 dataset, 64 estimators, descriptor set C, and shear modulus as target property.

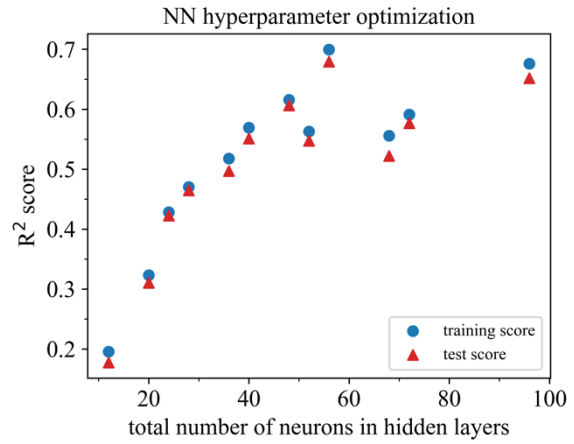

Figure S304: Hyperparameter optimization for NN trained with the V14 dataset, descriptor set AC, and shear modulus as target property.

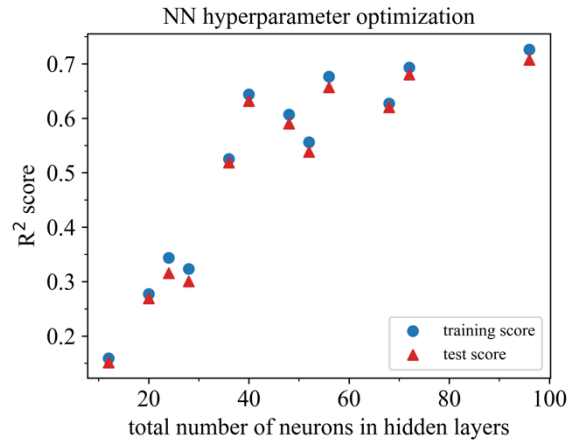

Figure S305: Hyperparameter optimization for NN trained with the V14 dataset, descriptor set A, and shear modulus as target property.

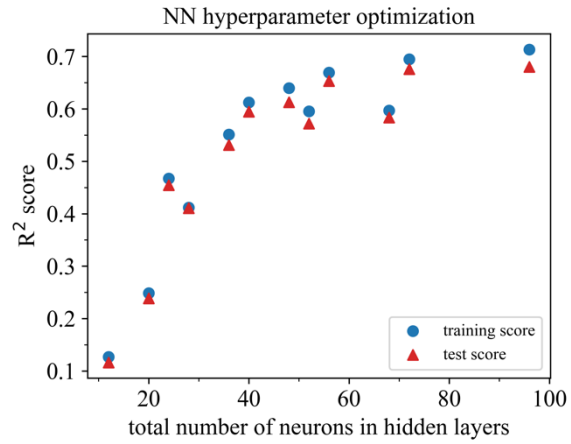

Figure S306: Hyperparameter optimization for NN trained with the V14 dataset, descriptor set C, and shear modulus as target property.

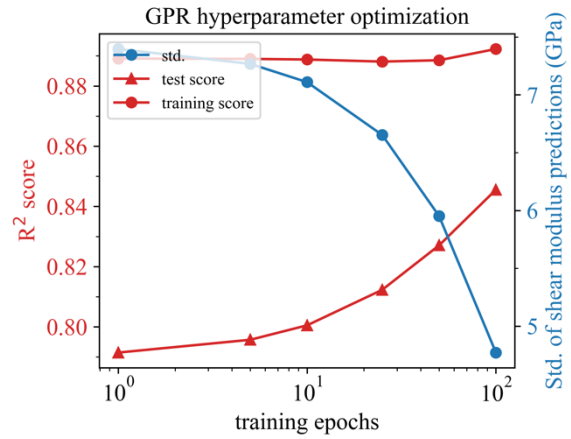

Figure S307: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.01, descriptor set AC, and shear modulus as target property.

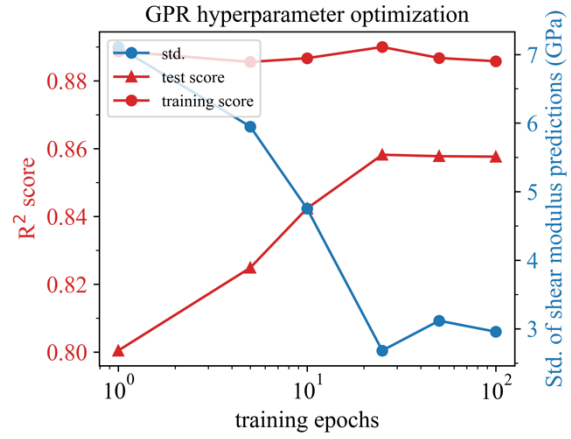

Figure S308: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.1, descriptor set AC, and shear modulus as target property.

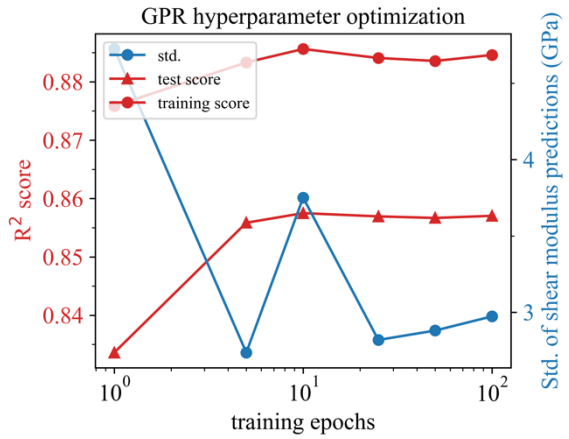

Figure S309: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 1.0, descriptor set AC, and shear modulus as target property.

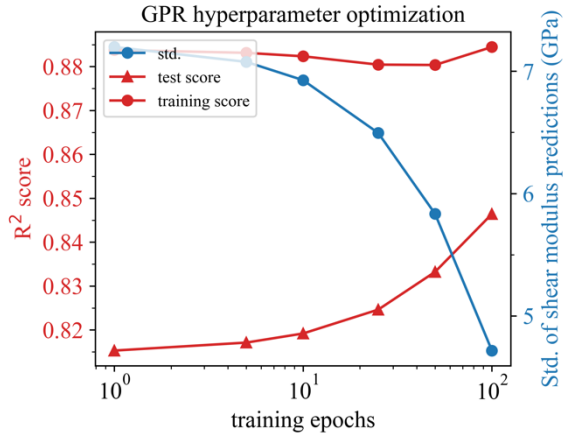

Figure S310: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.01, descriptor set A, and shear modulus as target property.

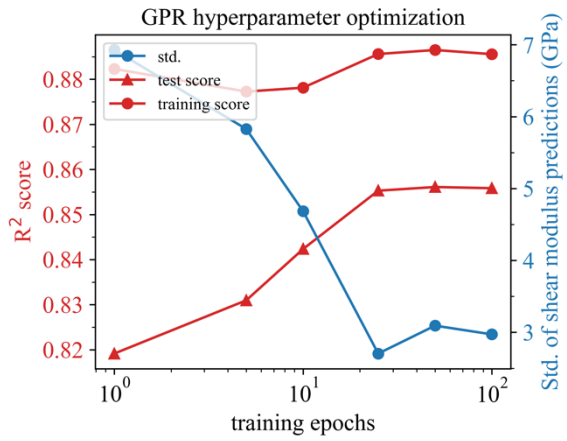

Figure S311: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.1, descriptor set A, and shear modulus as target property.

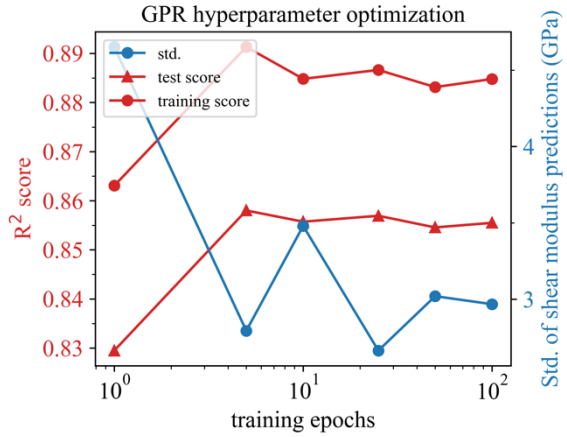

Figure S312: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 1.0, descriptor set A, and shear modulus as target property.

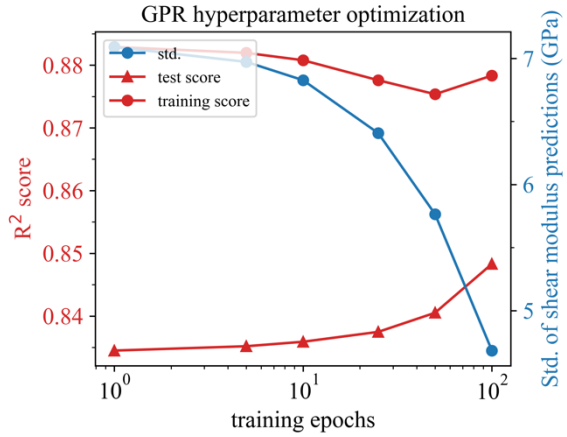

Figure S313: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.01, descriptor set C, and shear modulus as target property.

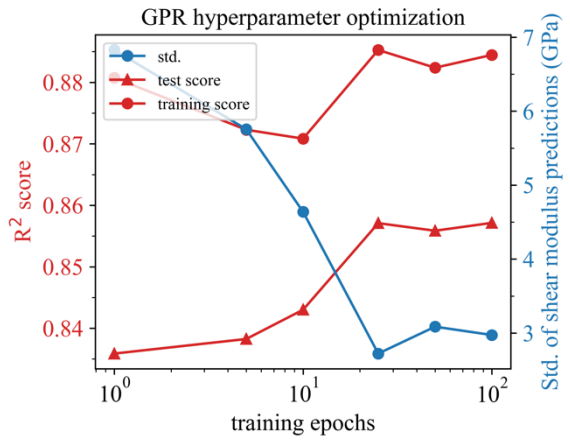

Figure S314: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.1, descriptor set C, and shear modulus as target property.

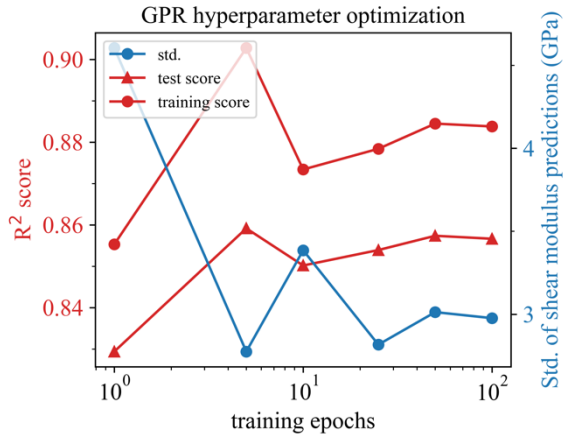

Figure S315: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 1.0, descriptor set C, and shear modulus as target property.

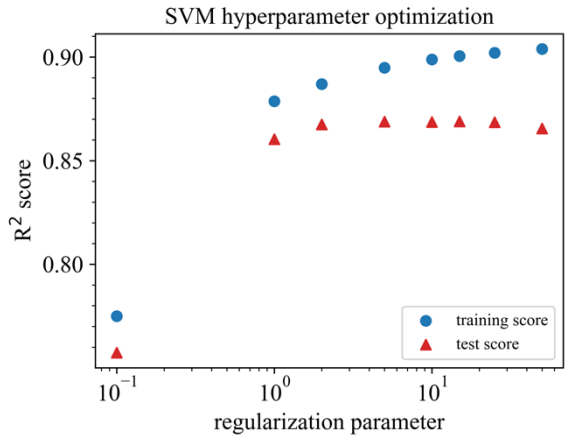

Figure S316: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.01$, descriptor set AC, and Young's modulus as target property.

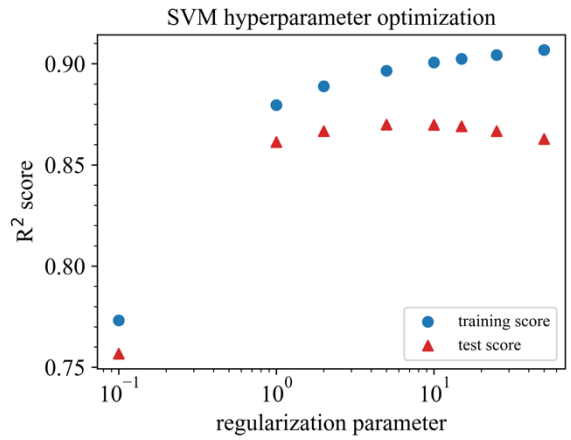

Figure S317: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.1$, descriptor set AC, and Young's modulus as target property.

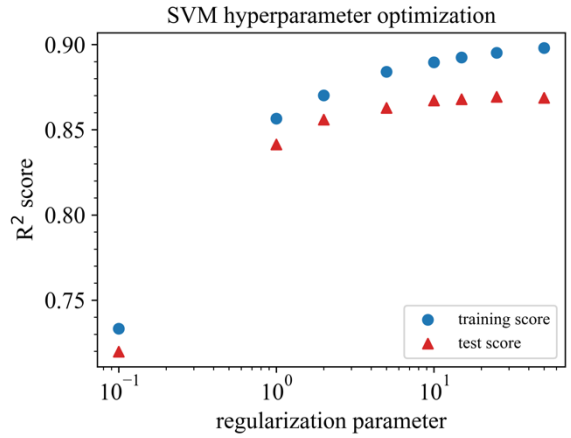

Figure S318: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.2$, descriptor set AC, and Young's modulus as target property.

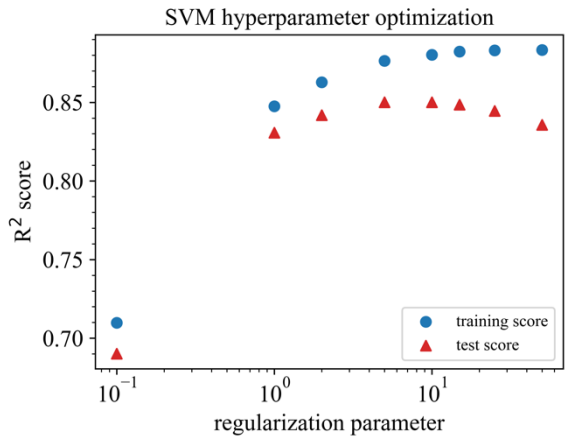

Figure S319: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.5$, descriptor set AC, and Young's modulus as target property.

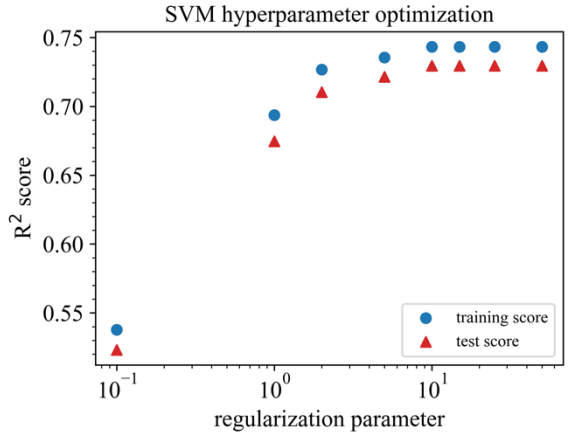

Figure S320: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 1.0$, descriptor set AC, and Young's modulus as target property.

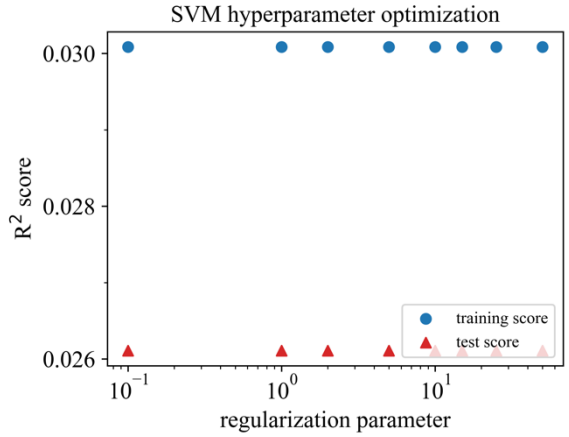

Figure S321: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 2.0$, descriptor set AC, and Young's modulus as target property.

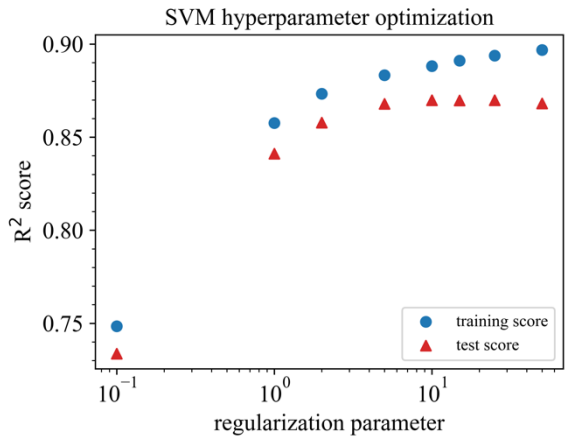

Figure S322: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.01$, descriptor set A, and Young's modulus as target property.

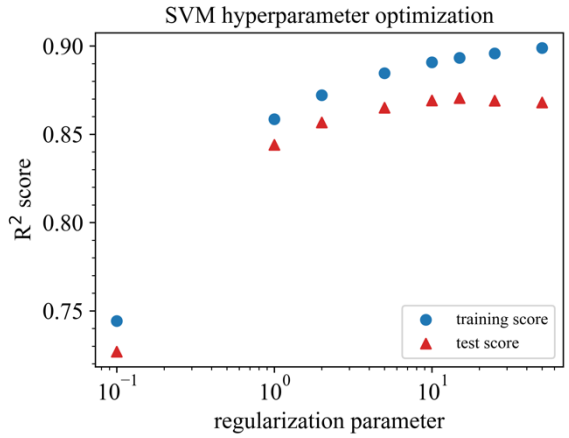

Figure S323: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.1$, descriptor set A, and Young's modulus as target property.

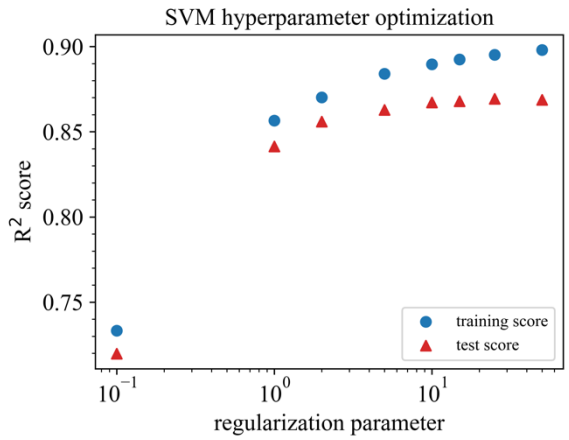

Figure S324: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.2$, descriptor set A, and Young's modulus as target property.

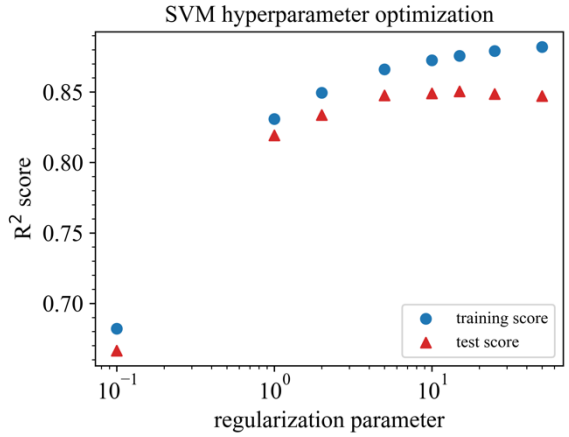

Figure S325: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.5$, descriptor set A, and Young's modulus as target property.

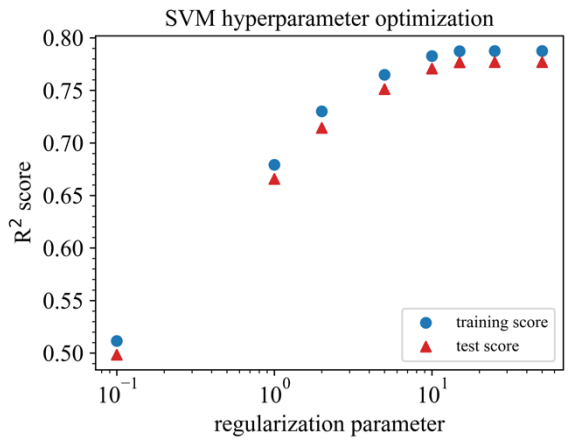

Figure S326: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 1.0$, descriptor set A, and Young's modulus as target property.

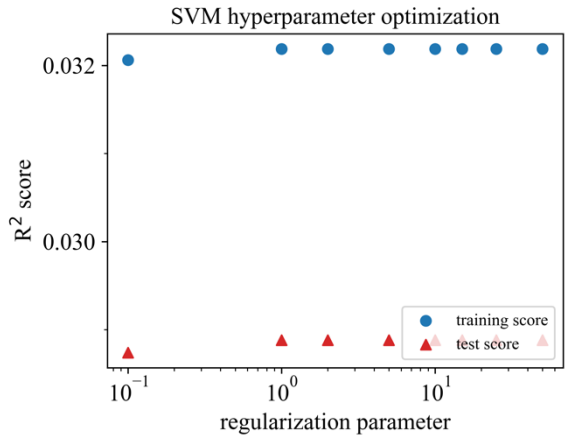

Figure S327: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 2.0$, descriptor set A, and Young's modulus as target property.

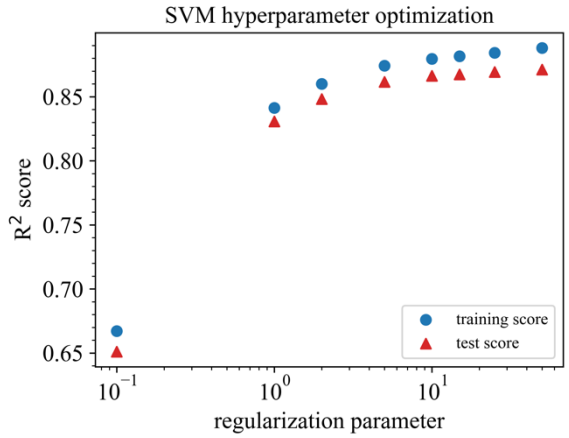

Figure S328: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.01$, descriptor set C, and Young's modulus as target property.

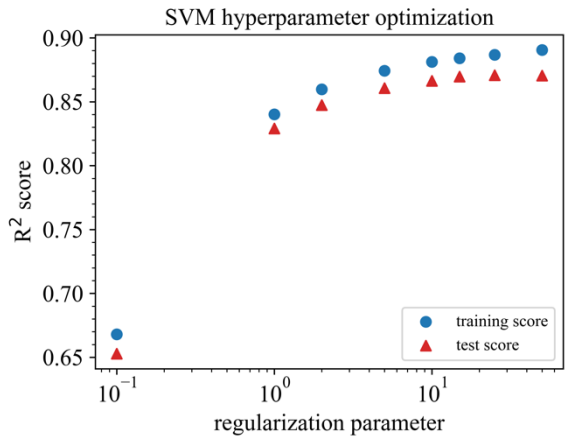

Figure S329: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.1$, descriptor set C, and Young's modulus as target property.

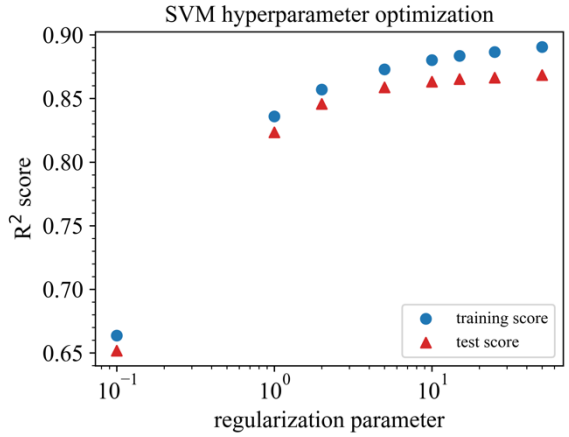

Figure S330: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.2$, descriptor set C, and Young's modulus as target property.

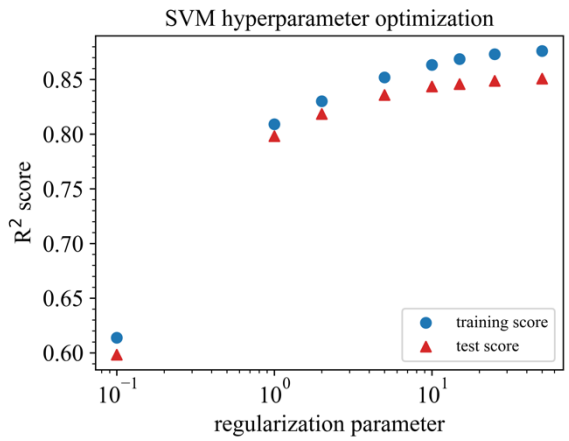

Figure S331: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, $\epsilon = 0.5$, descriptor set C, and Young's modulus as target property.

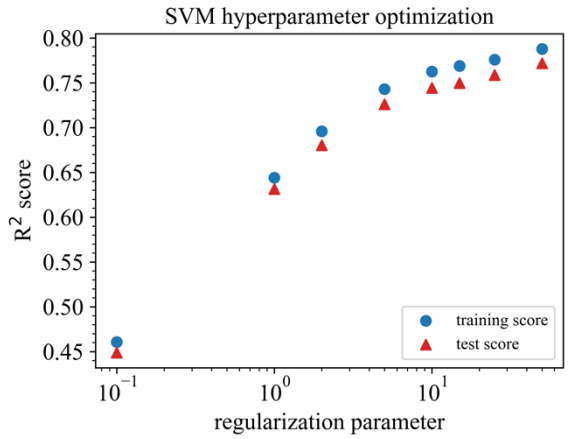

Figure S332: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, ϵ = 1.0, descriptor set C, and Young's modulus as target property.

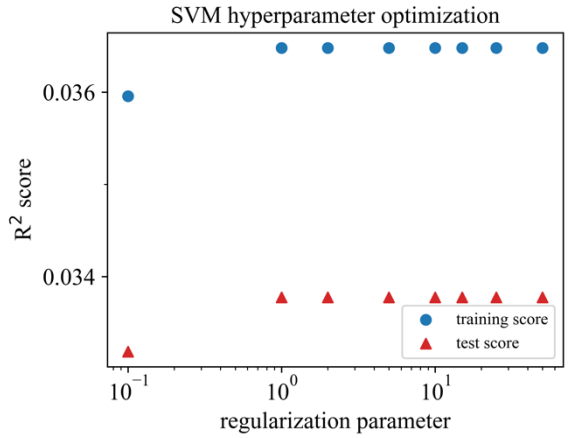

Figure S333: Hyperparameter optimization for SVM (SVR) trained with the V14 dataset, ϵ = 2.0, descriptor set C, and Young's modulus as target property.

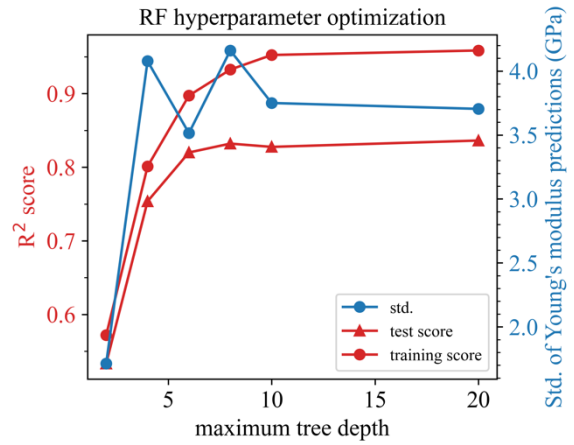

Figure S334: Hyperparameter optimization for RF trained with the V14 dataset, 4 estimators, descriptor set AC, and Young's modulus as target property.

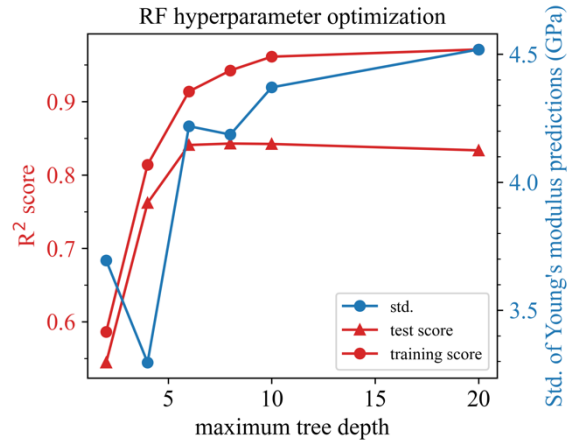

Figure S335: Hyperparameter optimization for RF trained with the V14 dataset, 8 estimators, descriptor set AC, and Young's modulus as target property.

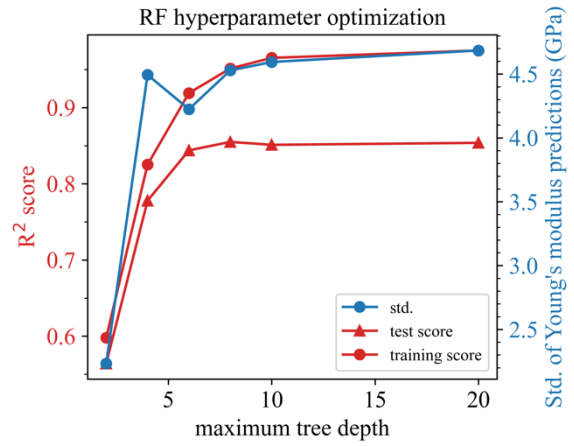

Figure S336: Hyperparameter optimization for RF trained with the V14 dataset, 16 estimators, descriptor set AC, and Young's modulus as target property.

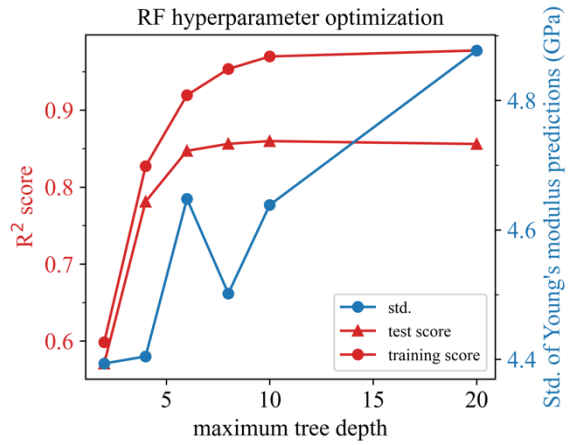

Figure S337: Hyperparameter optimization for RF trained with the V14 dataset, 32 estimators, descriptor set AC, and Young's modulus as target property.

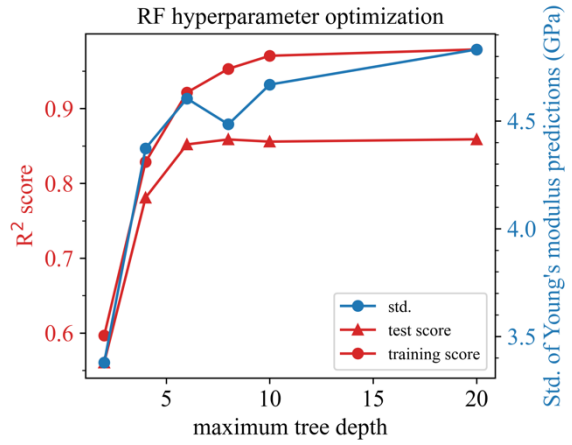

Figure S338: Hyperparameter optimization for RF trained with the V14 dataset, 64 estimators, descriptor set AC, and Young's modulus as target property.

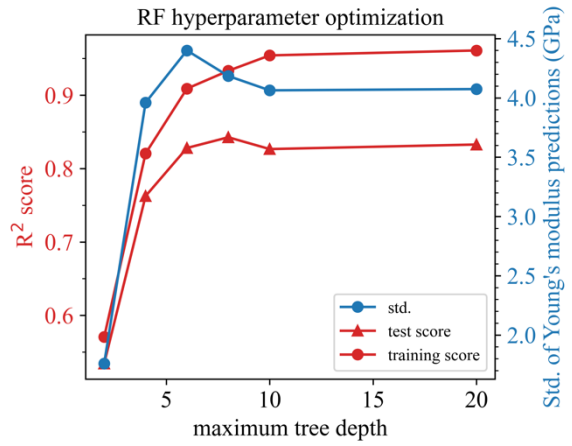

Figure S339: Hyperparameter optimization for RF trained with the V14 dataset, 4 estimators, descriptor set A, and Young's modulus as target property.

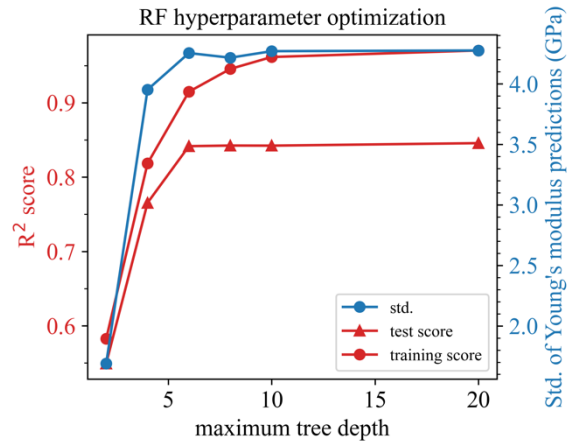

Figure S340: Hyperparameter optimization for RF trained with the V14 dataset, 8 estimators, descriptor set A, and Young's modulus as target property.

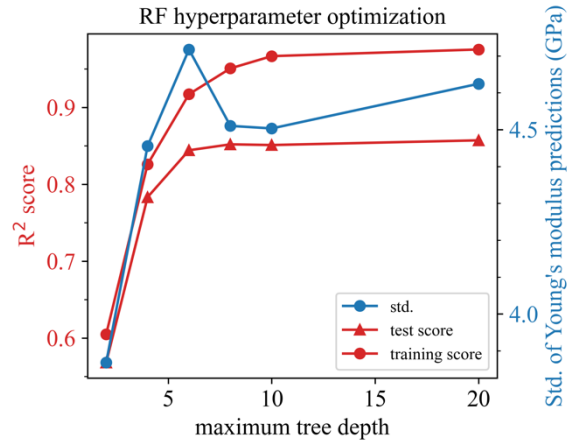

Figure S341: Hyperparameter optimization for RF trained with the V14 dataset, 16 estimators, descriptor set A, and Young's modulus as target property.

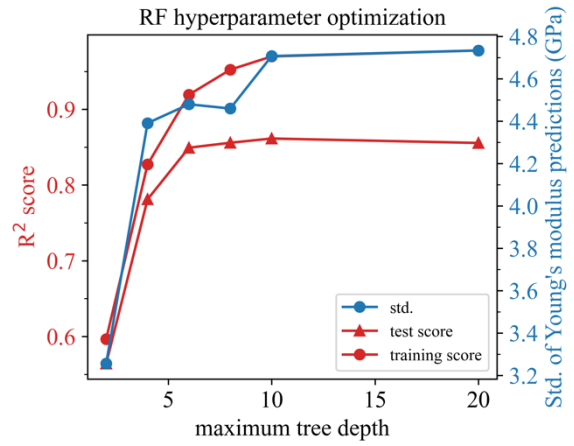

Figure S342: Hyperparameter optimization for RF trained with the V14 dataset, 32 estimators, descriptor set A, and Young's modulus as target property.

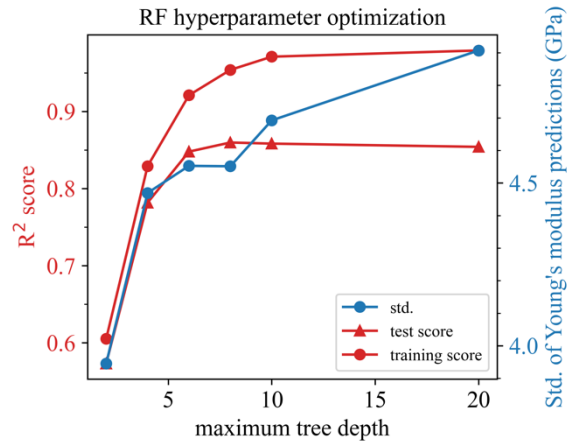

Figure S343: Hyperparameter optimization for RF trained with the V14 dataset, 64 estimators, descriptor set A, and Young's modulus as target property.

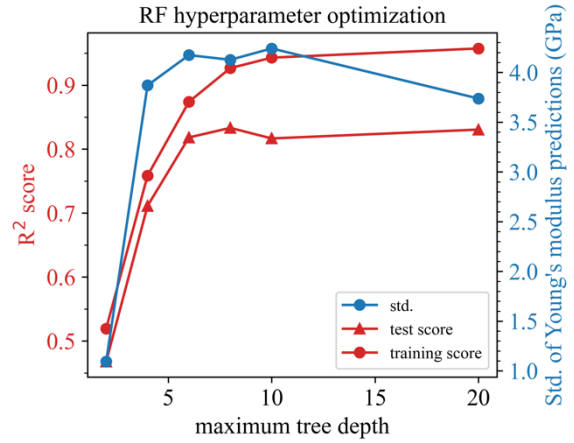

Figure S344: Hyperparameter optimization for RF trained with the V14 dataset, 4 estimators, descriptor set C, and Young's modulus as target property.

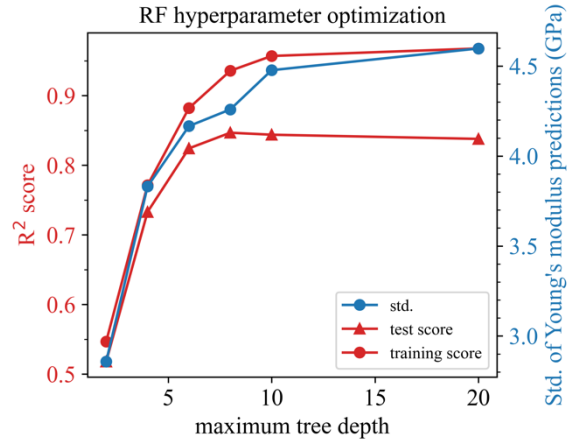

Figure S345: Hyperparameter optimization for RF trained with the V14 dataset, 8 estimators, descriptor set C, and Young's modulus as target property.

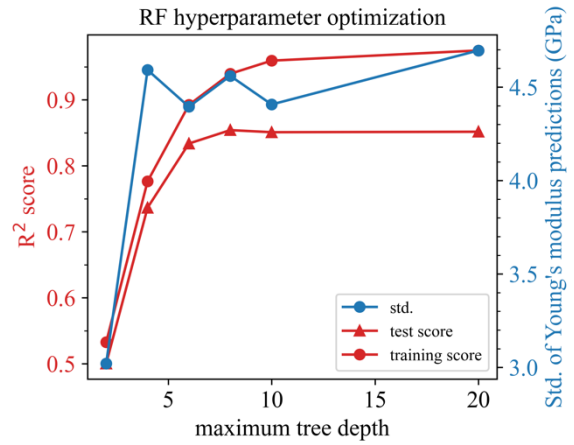

Figure S346: Hyperparameter optimization for RF trained with the V14 dataset, 16 estimators, descriptor set C, and Young's modulus as target property.

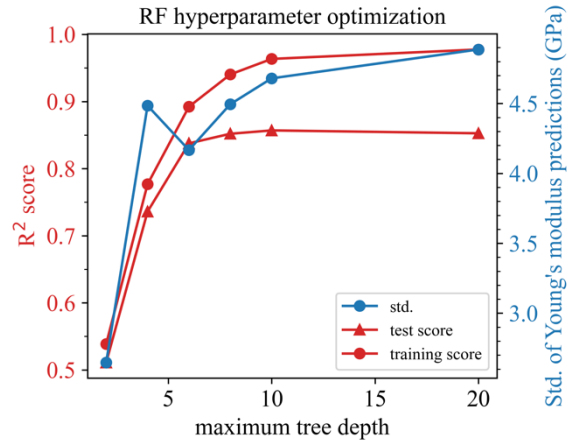

Figure S347: Hyperparameter optimization for RF trained with the V14 dataset, 32 estimators, descriptor set C, and Young's modulus as target property.

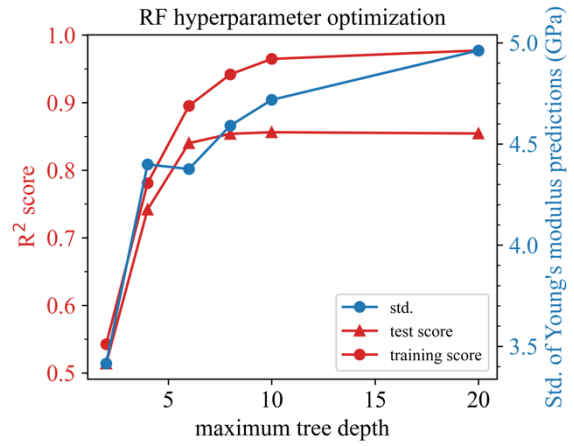

Figure S348: Hyperparameter optimization for RF trained with the V14 dataset, 64 estimators, descriptor set C, and Young's modulus as target property.

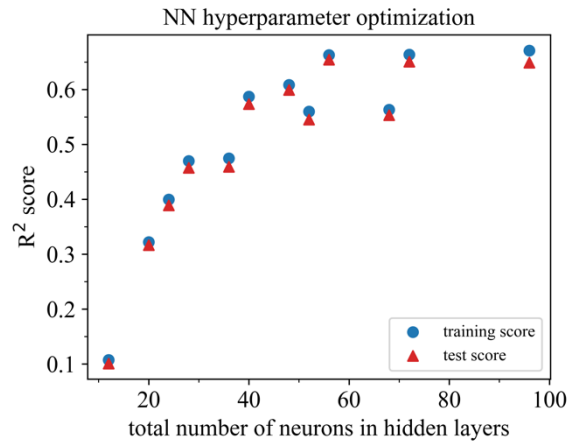

Figure S349: Hyperparameter optimization for NN trained with the V14 dataset, descriptor set AC, and Young's modulus as target property.

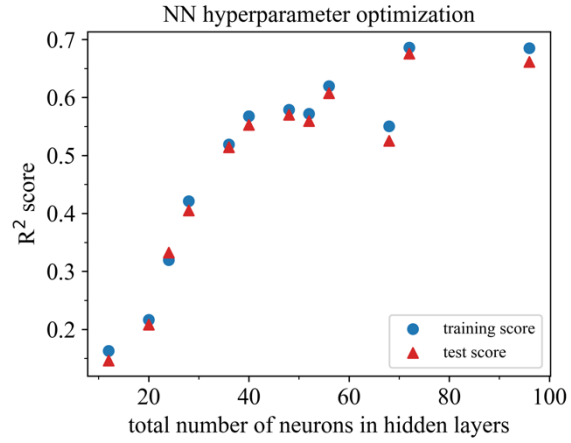

Figure S350: Hyperparameter optimization for NN trained with the V14 dataset, descriptor set A, and Young's modulus as target property.

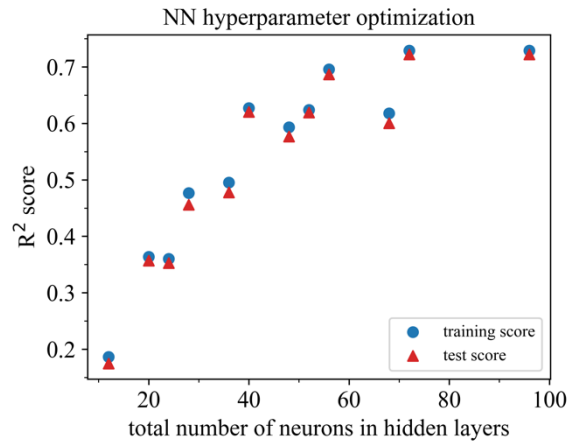

Figure S351: Hyperparameter optimization for NN trained with the V14 dataset, descriptor set C, and Young's modulus as target property.

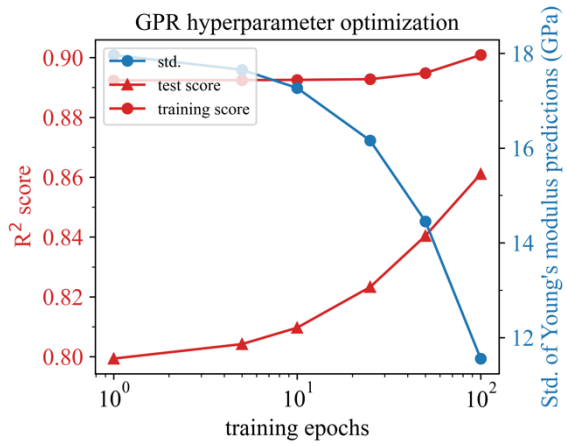

Figure S352: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.01, descriptor set AC, and Young's modulus as target property.

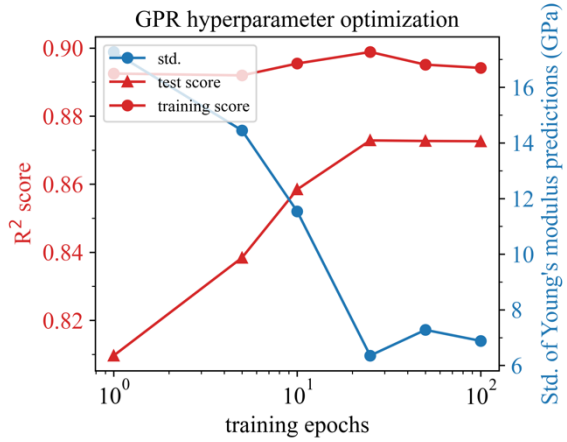

Figure S353: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.1, descriptor set AC, and Young's modulus as target property.

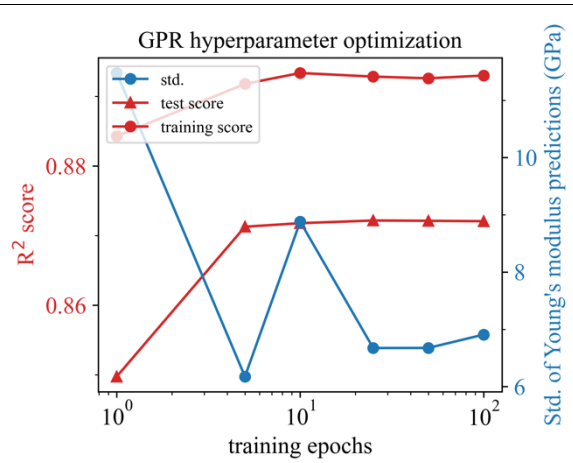

Figure S354: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 1.0, descriptor set AC, and Young's modulus as target property.

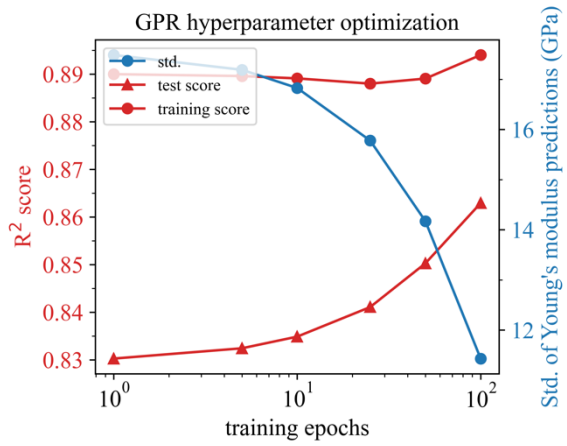

Figure S355: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.01, descriptor set A, and Young's modulus as target property.

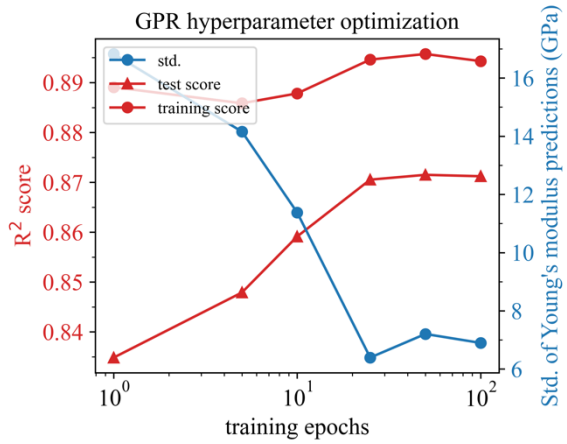

Figure S356: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.1, descriptor set A, and Young's modulus as target property.

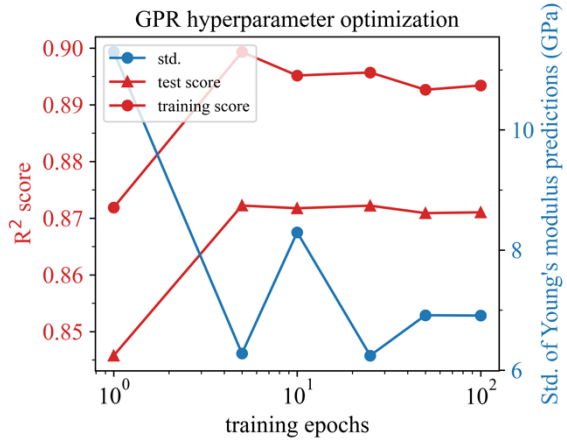

Figure S357: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 1.0, descriptor set A, and Young's modulus as target property.

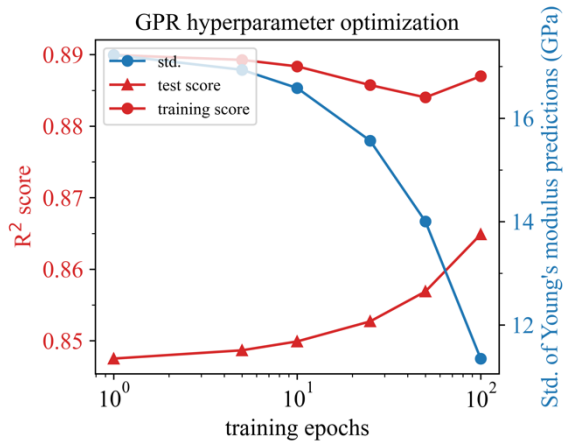

Figure S358: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.01, descriptor set C, and Young's modulus as target property.

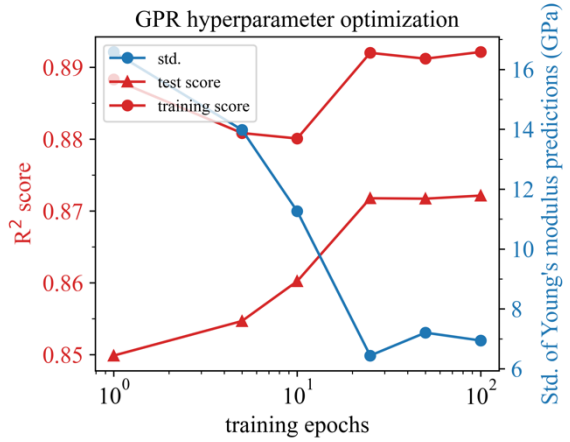

Figure S359: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 0.1, descriptor set C, and Young's modulus as target property.

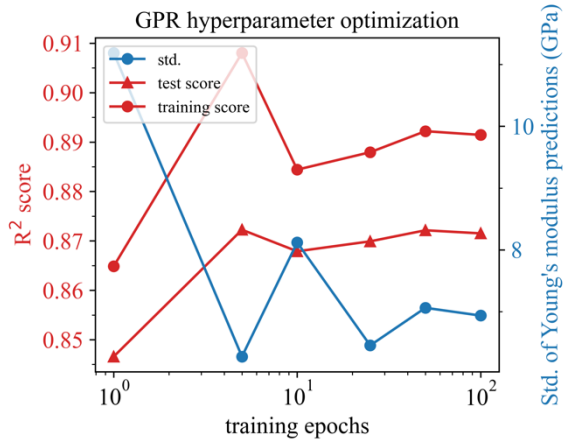

Figure S360: Hyperparameter optimization for GPR trained with the V14 dataset, a learning rate of 1.0, descriptor set C, and Young's modulus as target property.

Table S1: Selection of hyperparameter ε for SVM (SVR) trained with the V18 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

SVR	set AC	set A	set C
Density	0.2	0.2	0.2
Bulk modulus	0.2	0.2	0.2
Shear modulus	0.2	0.2	0.2
Young's modulus	0.2	0.2	0.2

Table S2: Selection of hyperparameter C for SVM (SVR) trained with the V18 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

SVR	set AC	set A	set C
Density	10	10	10
Bulk modulus	5	10	10
Shear modulus	10	10	10
Young's modulus	5	10	10

Table S3: Selection of hyperparameter for number of tree estimators for RF trained with the V18 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

RF	set AC	set A	set C
Density	16	16	16
Bulk modulus	16	16	16
Shear modulus	16	16	16
Young's modulus	16	16	16

Table S4: Selection of hyperparameter for maximum tree depth for RF trained with the V18 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

RF	set AC	set A	set C
Density	10	10	10
Bulk modulus	8	8	10
Shear modulus	8	8	8
Young's modulus	8	10	8

Table S5: Selection of NN structure represented the by total number of neurons for models trained with the V18 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

NN	set AC	set A	set C
Density	96	72	96
Bulk modulus	72	72	72
Shear modulus	96	56	56
Young's modulus	56	96	56

Table S6: Selection of hyperparameter for learning rate for GPR trained with the V18 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

GPR	set AC	set A	set C
Density	0.1	0.1	0.1
Bulk modulus	0.1	0.1	0.1
Shear modulus	0.1	0.1	0.1
Young's modulus	0.1	0.1	0.1

Table S7: Selection of hyperparameter for number of epochs for GPR trained with the V18 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

GPR	set AC	set A	set C
Density	25	25	25
Bulk modulus	25	25	25
Shear modulus	25	25	25
Young's modulus	25	25	25

Table S8: Selection of hyperparameter ε for SVM (SVR) trained with the V14 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

SVR	set AC	set A	set C
Density	0.2	0.1	0.1
Bulk modulus	0.2	0.2	0.2
Shear modulus	0.2	0.2	0.2
Young's modulus	0.2	0.2	0.1

Table S9: Selection of hyperparameter C for SVM (SVR) trained with the V14 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

SVR	set AC	set A	set C
Density	5	5	10
Bulk modulus	10	10	10
Shear modulus	5	5	10
Young's modulus	5	5	10

Table S10: Selection of hyperparameter for number of tree estimators for RF trained with the V14 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

RF	set AC	set A	set C
Density	16	16	16
Bulk modulus	16	16	16
Shear modulus	16	16	16
Young's modulus	16	16	16

Table S11: Selection of hyperparameter for maximum tree depth for RF trained with the V14 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

RF	set AC	set A	set C
Density	8	10	10
Bulk modulus	10	10	10
Shear modulus	8	8	8
Young's modulus	8	8	8

Table S12: Selection of NN structure represented the by total number of neurons for models trained with the V14 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

NN	set AC	set A	set C
Density	56	72	72
Bulk modulus	40	56	48
Shear modulus	56	56	56
Young's modulus	56	72	72

Table S13: Selection of hyperparameter for learning rate for GPR trained with the V14 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

GPR	set AC	set A	set C

Density	0.1	0.1	0.1
Bulk modulus	0.1	0.1	0.1
Shear modulus	0.1	0.1	0.1
Young's modulus	0.1	0.1	0.1

Table S14: Selection of hyperparameter for number of epochs for GPR trained with the V14 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

GPR	set AC	set A	set C
Density	50	50	50
Bulk modulus	25	25	25
Shear modulus	25	25	25
Young's modulus	25	25	25

Table S15: R^2 scores for different ML algorithms and descriptor sets for predicting glass density for the V18 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

Density	set AC	set A	set C
SVR	0.94	0.94	0.94
NN	0.84	0.85	0.87
RF	0.98	0.97	0.98
GPR	0.95	0.95	0.95

Table S16: R^2 scores for different ML algorithms and descriptor sets for predicting bulk modulus for the V18 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

Bulk modulus	set AC	set A	set C
SVR	0.91	0.91	0.91
NN	0.69	0.73	0.73
RF	0.91	0.91	0.92
GPR	0.91	0.91	0.92

Table S17: R^2 scores for different ML algorithms and descriptor sets for predicting shear modulus for the V18 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

Shear modulus	set AC	set A	set C
SVR	0.89	0.88	0.89
NN	0.74	0.77	0.72
RF	0.90	0.90	0.89
GPR	0.90	0.89	0.90

Table S18: R^2 scores for different ML algorithms and descriptor sets for predicting Young's modulus for the V18 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

Young's modulus	set AC	set A	set C
SVR	0.90	0.90	0.90
NN	0.72	0.69	0.74
RF	0.91	0.91	0.90
GPR	0.91	0.91	0.91

Table S19: R^2 scores for different ML algorithms and descriptor sets for predicting glass density for the V14 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

Density	set AC	set A	set C
SVR	0.98	0.97	0.97
NN	0.78	0.81	0.82
RF	0.96	0.96	0.97
GPR	0.98	0.97	0.98

Table S20: R^2 scores for different ML algorithms and descriptor sets for predicting bulk modulus for the V14 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

Bulk modulus	set AC	set A	set C
SVR	0.86	0.86	0.85
NN	0.56	0.58	0.62
RF	0.83	0.84	0.84
GPR	0.86	0.86	0.86

Table S21: R^2 scores for different ML algorithms and descriptor sets for predicting shear modulus for the V14 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

Shear modulus	set AC	set A	set C
SVR	0.86	0.85	0.85
NN	0.64	0.64	0.63
RF	0.84	0.84	0.83
GPR	0.86	0.86	0.86

Table S22: R^2 scores for different ML algorithms and descriptor sets for predicting Young's modulus for the V14 dataset (set C: compositional descriptors, set A: ab initio derived descriptors, set AC: combination of both).

Young's modulus	set AC	set A	set C
SVR	0.87	0.86	0.87
NN	0.64	0.69	0.64
RF	0.85	0.85	0.85
GPR	0.87	0.87	0.87

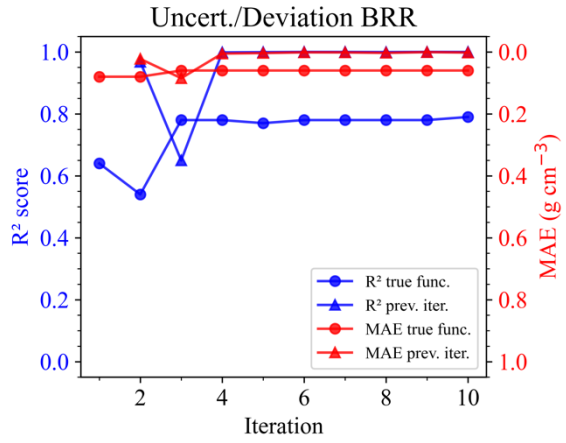

Figure S361: Evaluation of a compositional selection method (by the highest discrepancy between $\hat{\rho}_{\mathrm{MD}}(\mathbf{x})$ and $\hat{\rho}_{\mathrm{SCALE/SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using BRR.

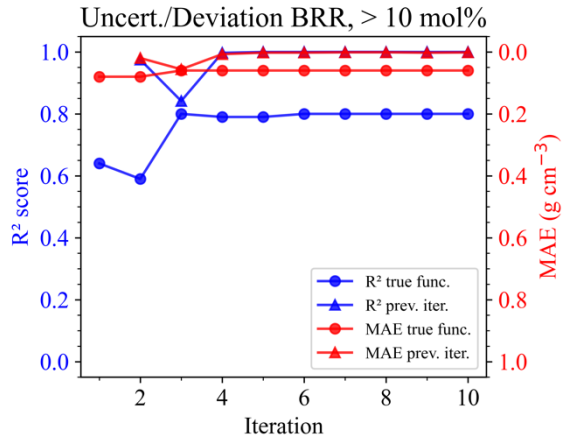

Figure S362: Evaluation of a compositional selection method (by the highest discrepancy between $\hat{\rho}_{\mathrm{MD}}(\mathbf{x})$ and $\hat{\rho}_{\mathrm{SCALE/SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using BRR. A minimum distance criterion between compositions of 10 mol% is applied.

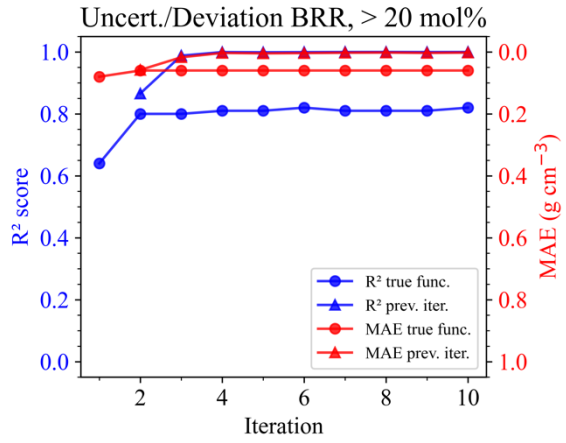

Figure S363: Evaluation of a compositional selection method (by the highest discrepancy between $\hat{\rho}_{\mathrm{MD}}(\mathbf{x})$ and $\hat{\rho}_{\mathrm{SCALE/SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using BRR. A minimum distance criterion between compositions of 20 mol% is applied.

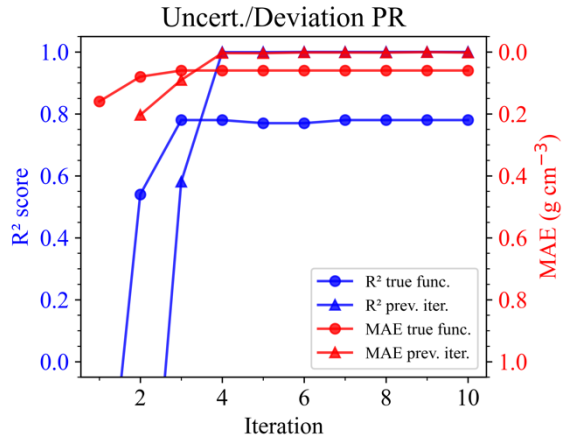

Figure S364: Evaluation of a compositional selection method (by the highest discrepancy between $\hat{\rho}_{\mathrm{MD}}(\mathbf{x})$ and $\hat{\rho}_{\mathrm{SCALE/SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using PR.

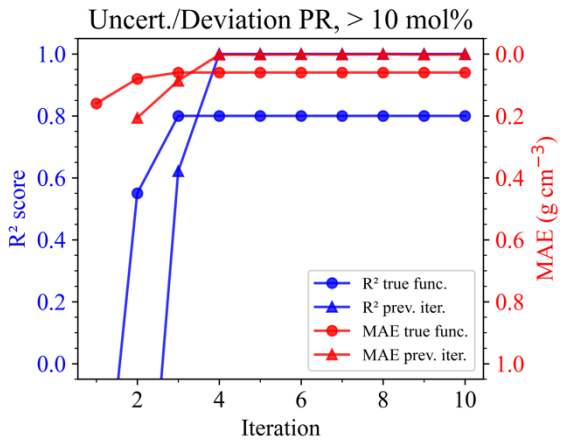

Figure S365: Evaluation of a compositional selection method (by the highest discrepancy between $\hat{\rho}_{\mathrm{MD}}(\mathbf{x})$ and $\hat{\rho}_{\mathrm{SCALE/SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using PR. A minimum distance criterion between compositions of 10 mol% is applied.

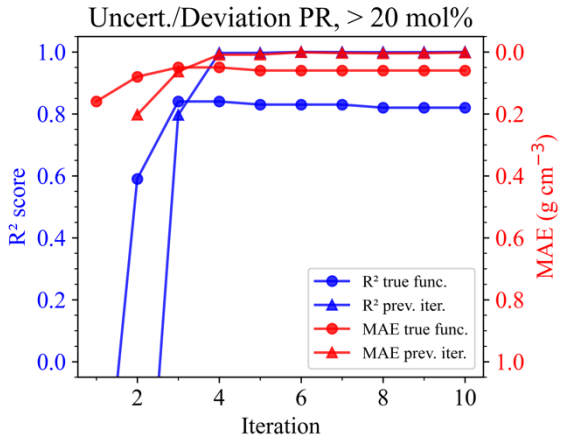

Figure S366: Evaluation of a compositional selection method (by the highest discrepancy between $\hat{\rho}_{\mathrm{MD}}(\mathbf{x})$ and $\hat{\rho}_{\mathrm{SCALE/SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using PR. A minimum distance criterion between compositions of 20 mol% is applied.

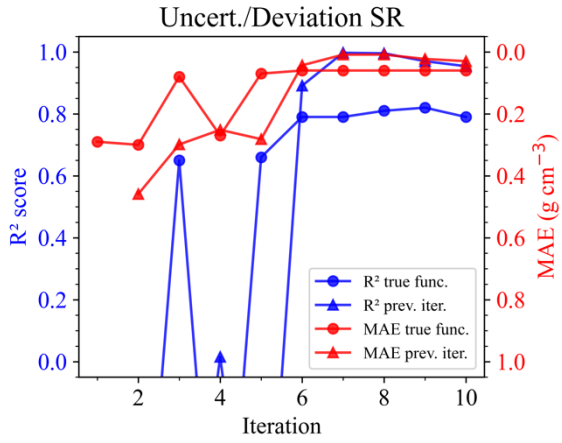

Figure S367: Evaluation of a compositional selection method (by the highest discrepancy between $\hat{\rho}_{\mathrm{MD}}(\mathbf{x})$ and $\hat{\rho}_{\mathrm{SCALE/SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using SR.

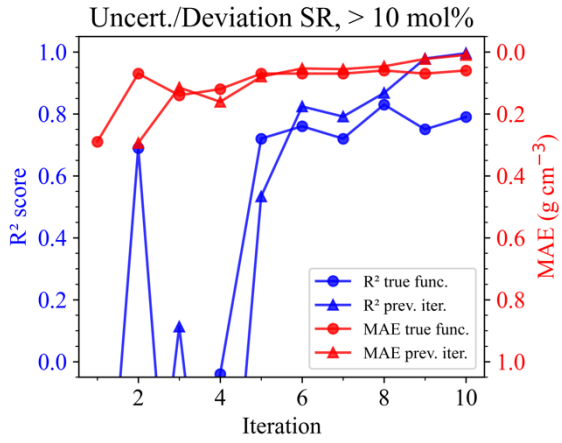

Figure S368: Evaluation of a compositional selection method (by the highest discrepancy between $\hat{\rho}_{\mathrm{MD}}(\mathbf{x})$ and $\hat{\rho}_{\mathrm{SCALE/SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using SR. A minimum distance criterion between compositions of 10 mol% is applied.

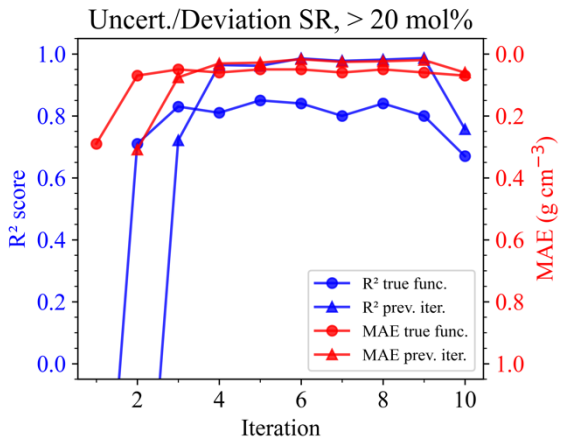

Figure S369: Evaluation of a compositional selection method (by the highest discrepancy between $\hat{\rho}_{\mathrm{MD}}(\mathbf{x})$ and $\hat{\rho}_{\mathrm{SCALE/SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using SR. A minimum distance criterion between compositions of 20 mol% is applied.

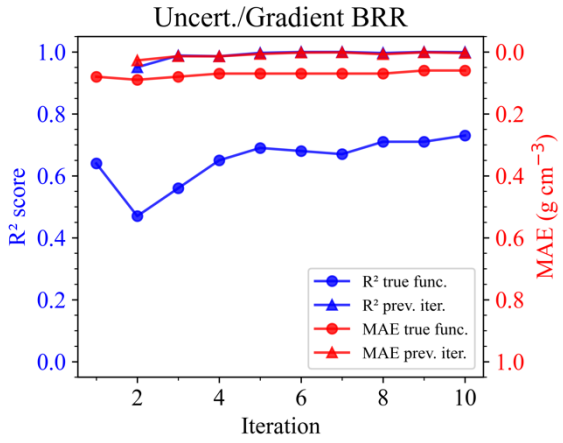

Figure S370: Evaluation of a compositional selection method (by the highest gradient of $\hat{\rho}_{\mathrm{SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using BRR.

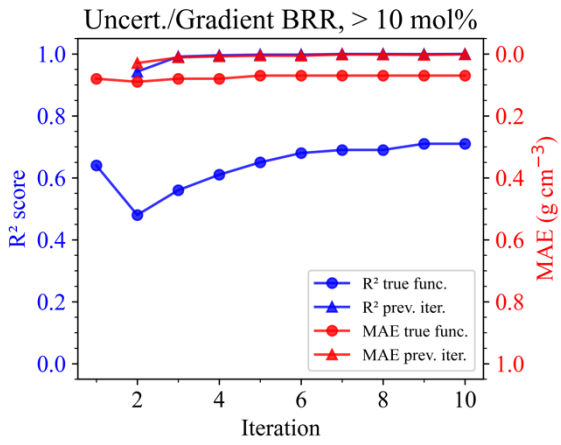

Figure S371: Evaluation of a compositional selection method (by the highest gradient of $\hat{\rho}_{\mathrm{SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using BRR. A minimum distance criterion between compositions of 10 mol% is applied.

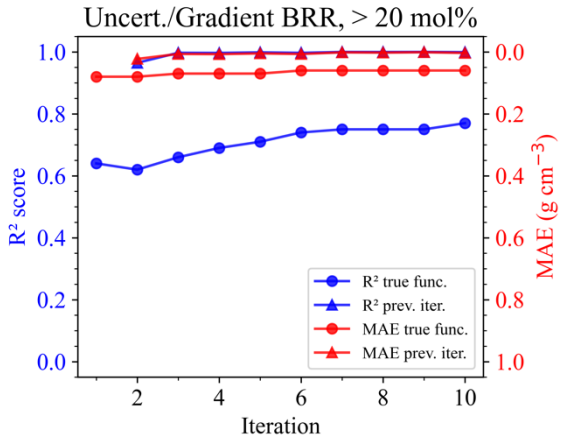

Figure S372: Evaluation of a compositional selection method (by the highest gradient of $\hat{\rho}_{\mathrm{SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using BRR. A minimum distance criterion between compositions of 20 mol% is applied.

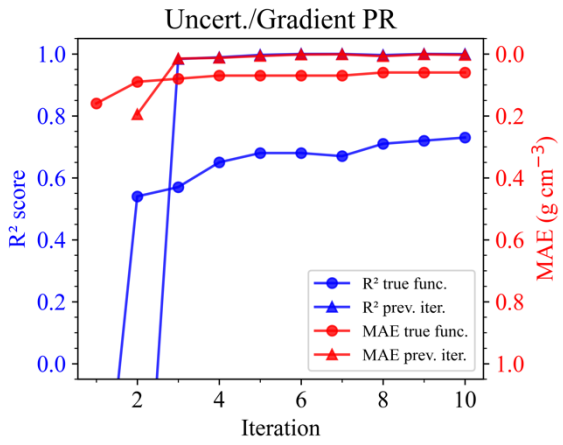

Figure S373: Evaluation of a compositional selection method (by the highest gradient of $\hat{\rho}_{\mathrm{SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using PR.

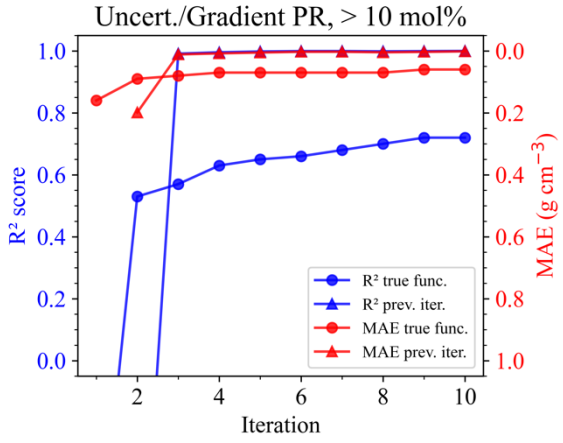

Figure S374: Evaluation of a compositional selection method (by the highest gradient of $\hat{\rho}_{\mathrm{SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using PR. A minimum distance criterion between compositions of 10 mol% is applied.

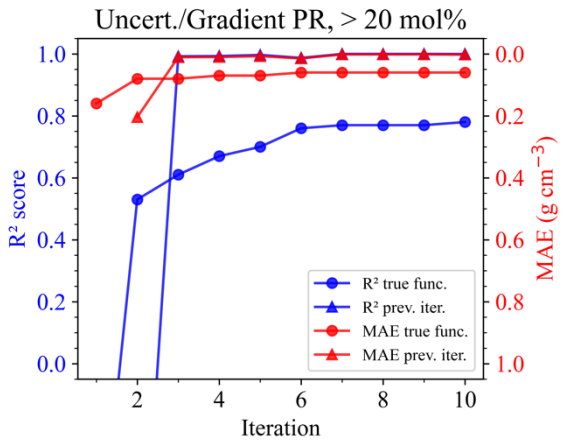

Figure S375: Evaluation of a compositional selection method (by the highest gradient of $\hat{\rho}_{\mathrm{SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using PR. A minimum distance criterion between compositions of 20 mol% is applied.

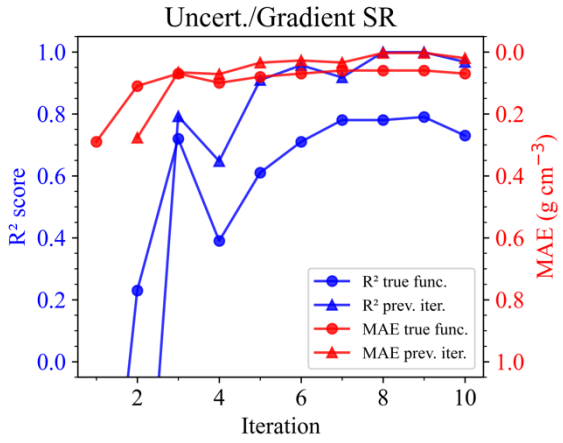

Figure S376: Evaluation of a compositional selection method (by the highest gradient of $\hat{\rho}_{\mathrm{SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using SR.

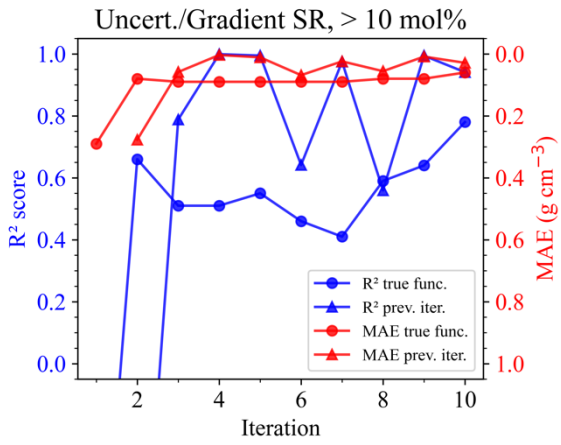

Figure S377: Evaluation of a compositional selection method (by the highest gradient of $\hat{\rho}_{\mathrm{SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using SR. A minimum distance criterion between compositions of 10 mol% is applied.

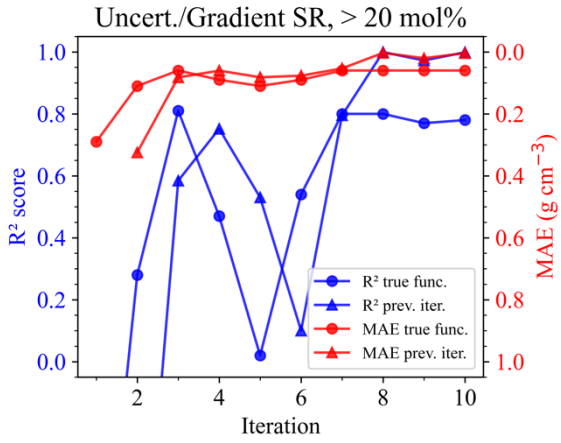

Figure S378: Evaluation of a compositional selection method (by the highest gradient of $\hat{\rho}_{\mathrm{SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using SR. A minimum distance criterion between compositions of 20 mol% is applied.

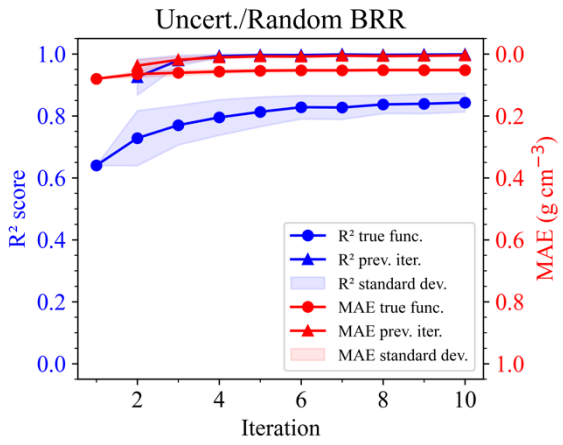

Figure S379: Evaluation of a compositional selection method (random selection) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using BRR.

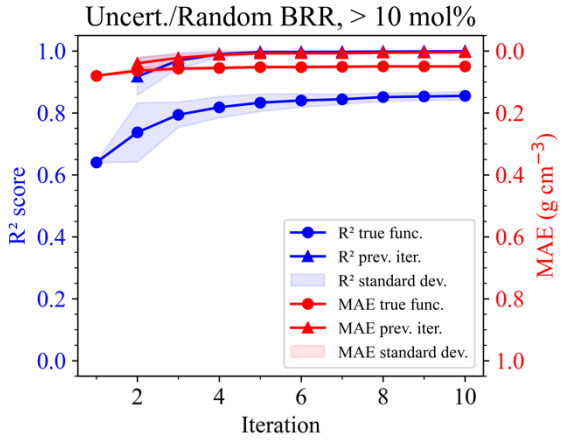

Figure S380: Evaluation of a compositional selection method (random selection) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using BRR. A minimum distance criterion between compositions of 10 mol% is applied.

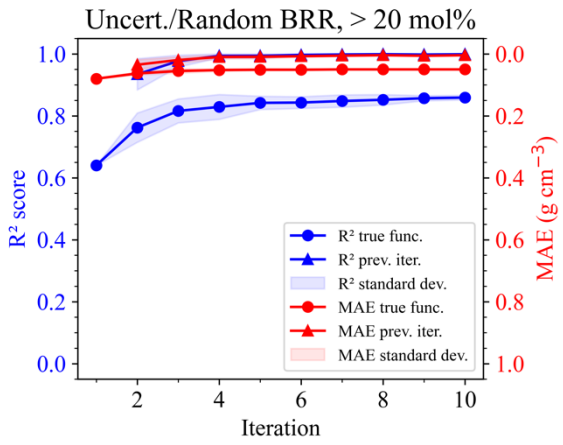

Figure S381: Evaluation of a compositional selection method (random selection) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using BRR. A minimum distance criterion between compositions of 20 mol% is applied.

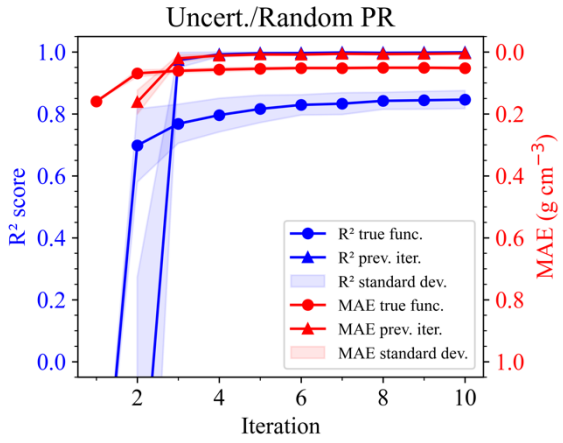

Figure S382: Evaluation of a compositional selection method (random selection) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using PR.

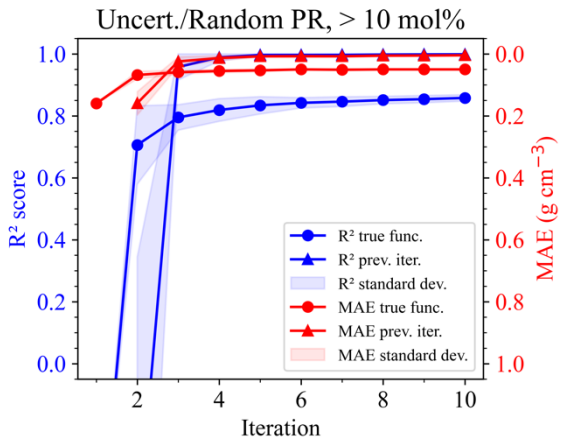

Figure S383: Evaluation of a compositional selection method (random selection) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using PR. A minimum distance criterion between compositions of 10 mol% is applied.

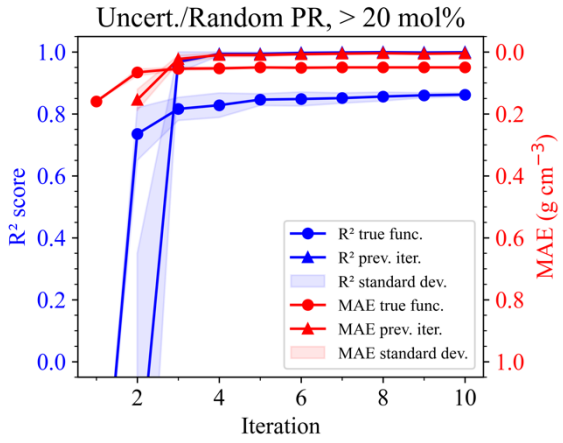

Figure S384: Evaluation of a compositional selection method (random selection) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using PR. A minimum distance criterion between compositions of 20 mol% is applied.

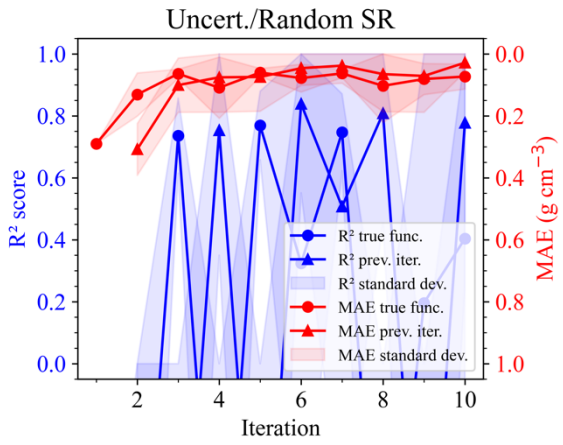

Figure S385: Evaluation of a compositional selection method (random selection) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using SR.

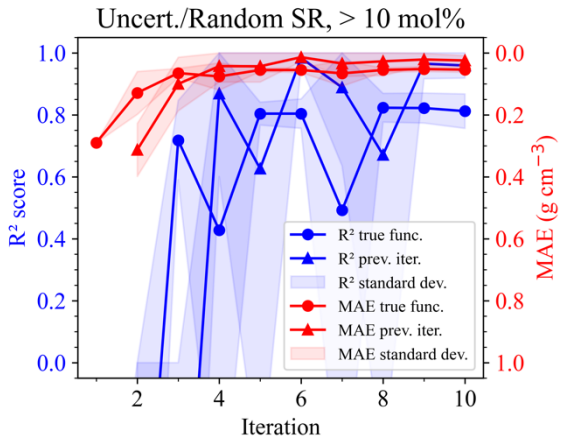

Figure S386: Evaluation of a compositional selection method (random selection) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using SR. A minimum distance criterion between compositions of 10 mol% is applied.

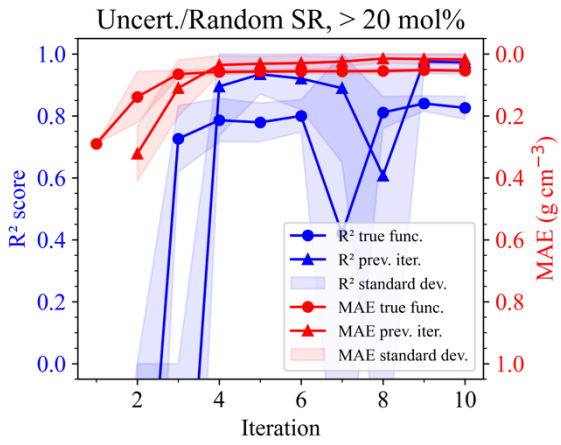

Figure S387: Evaluation of a compositional selection method (random selection) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is based on uncertainty, the models are determined using SR. A minimum distance criterion between compositions of 20 mol% is applied.

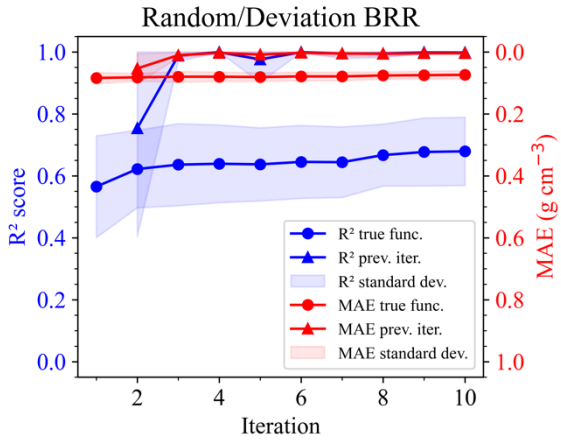

Figure S388: Evaluation of a compositional selection method (by the highest discrepancy between $\hat{\rho}_{\mathrm{MD}}(\mathbf{x})$ and $\hat{\rho}_{\mathrm{SCALE/SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using BRR.

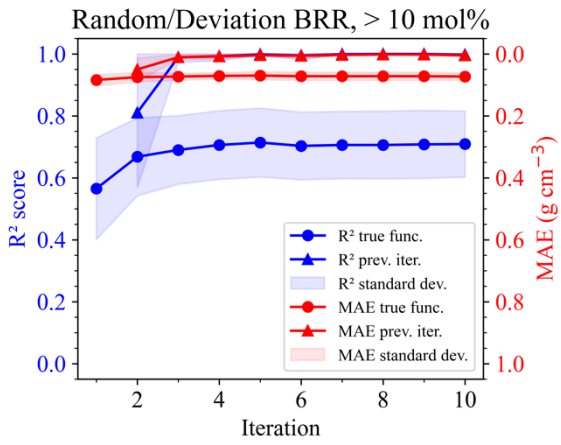

Figure S389: Evaluation of a compositional selection method (by the highest discrepancy between $\hat{\rho}_{\mathrm{MD}}(\mathbf{x})$ and $\hat{\rho}_{\mathrm{SCALE/SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using BRR. A minimum distance criterion between compositions of 10 mol% is applied.

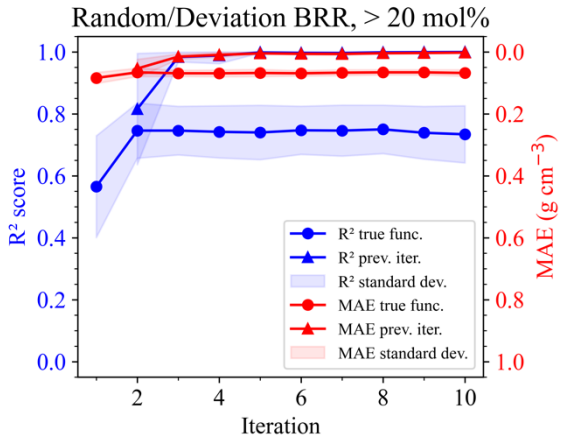

Figure S390: Evaluation of a compositional selection method (by the highest discrepancy between $\hat{\rho}_{\mathrm{MD}}(\mathbf{x})$ and $\hat{\rho}_{\mathrm{SCALE/SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using BRR. A minimum distance criterion between compositions of 20 mol% is applied.

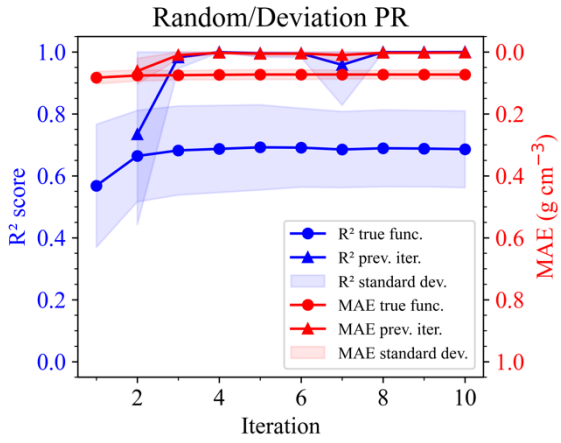

Figure S391: Evaluation of a compositional selection method (by the highest discrepancy between $\hat{\rho}_{\mathrm{MD}}(\mathbf{x})$ and $\hat{\rho}_{\mathrm{SCALE/SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using PR.

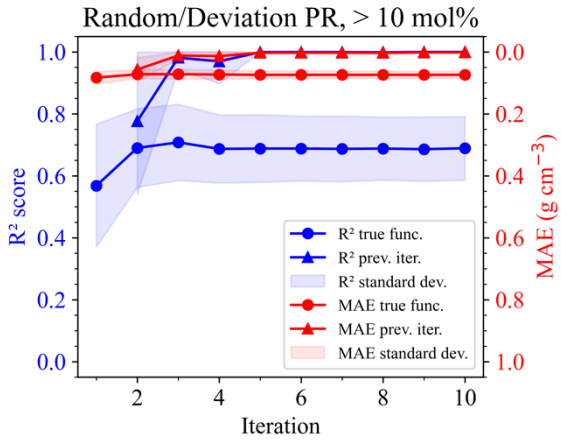

Figure S392: Evaluation of a compositional selection method (by the highest discrepancy between $\hat{\rho}_{\mathrm{MD}}(\mathbf{x})$ and $\hat{\rho}_{\mathrm{SCALE/SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using PR. A minimum distance criterion between compositions of 10 mol% is applied.

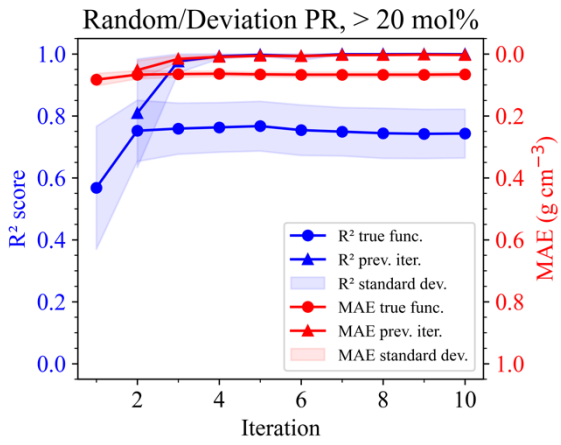

Figure S393: Evaluation of a compositional selection method (by the highest discrepancy between $\hat{\rho}_{\mathrm{MD}}(\mathbf{x})$ and $\hat{\rho}_{\mathrm{SCALE/SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using PR. A minimum distance criterion between compositions of 20 mol% is applied.

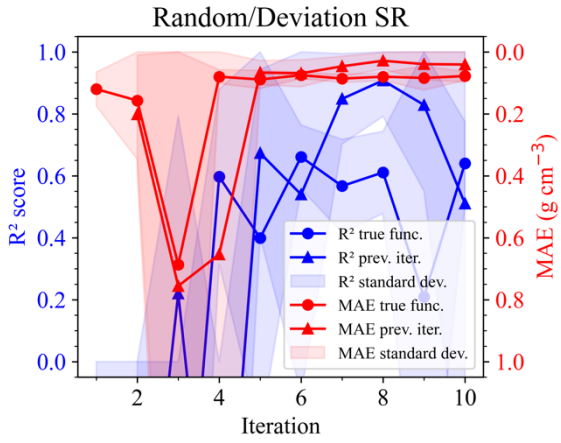

Figure S394: Evaluation of a compositional selection method (by the highest discrepancy between $\hat{\rho}_{\mathrm{MD}}(\mathbf{x})$ and $\hat{\rho}_{\mathrm{SCALE/SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using SR.

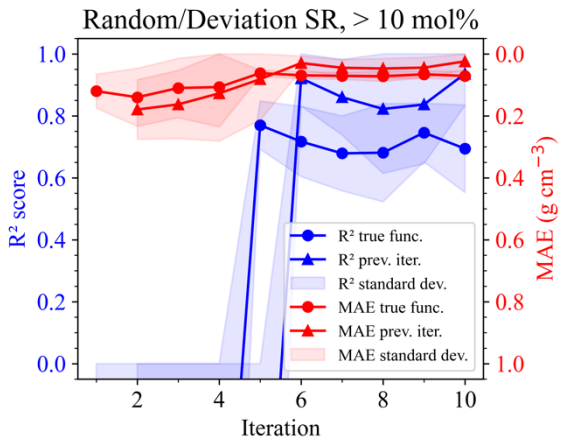

Figure S395: Evaluation of a compositional selection method (by the highest discrepancy between $\hat{\rho}_{\mathrm{MD}}(\mathbf{x})$ and $\hat{\rho}_{\mathrm{SCALE/SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using SR. A minimum distance criterion between compositions of 10 mol% is applied.

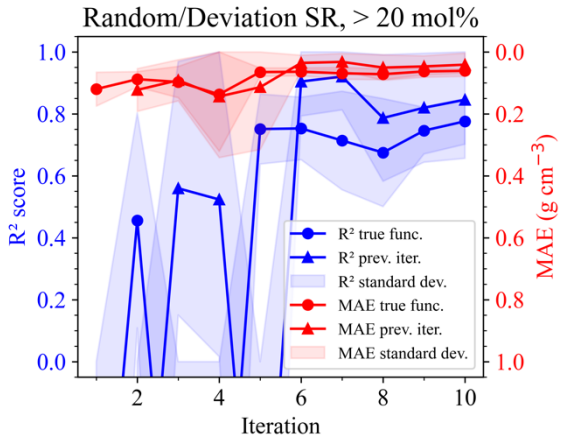

Figure S396: Evaluation of a compositional selection method (by the highest discrepancy between $\hat{\rho}_{\mathrm{MD}}(\mathbf{x})$ and $\hat{\rho}_{\mathrm{SCALE/SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using SR. A minimum distance criterion between compositions of 20 mol% is applied.

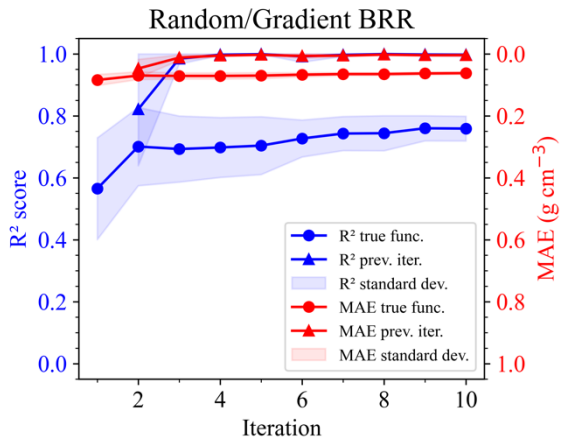

Figure S397: Evaluation of a compositional selection method (by the highest gradient of $\hat{\rho}_{\mathrm{SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using BRR.

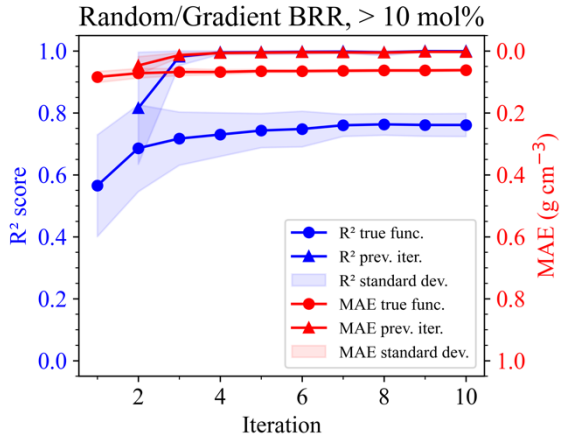

Figure S398: Evaluation of a compositional selection method (by the highest gradient of $\hat{\rho}_{\mathrm{SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using BRR. A minimum distance criterion between compositions of 10 mol% is applied.

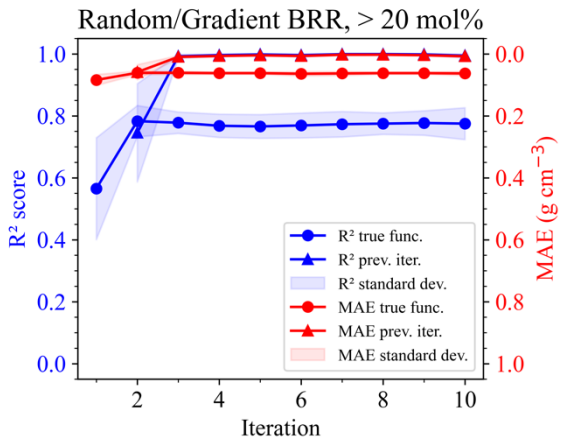

Figure S399: Evaluation of a compositional selection method (by the highest gradient of $\hat{\rho}_{\mathrm{SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using BRR. A minimum distance criterion between compositions of 20 mol% is applied.

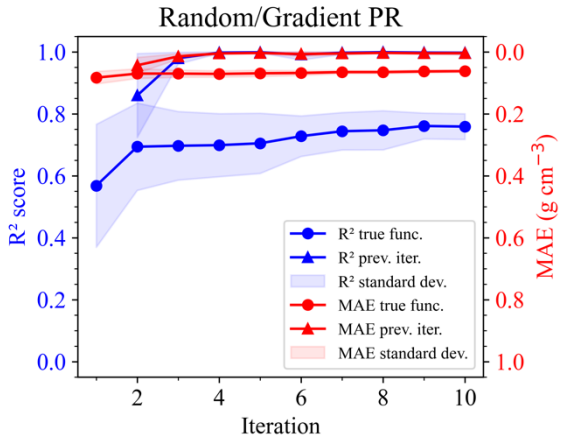

Figure S400: Evaluation of a compositional selection method (by the highest gradient of $\hat{\rho}_{\mathrm{SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using PR.

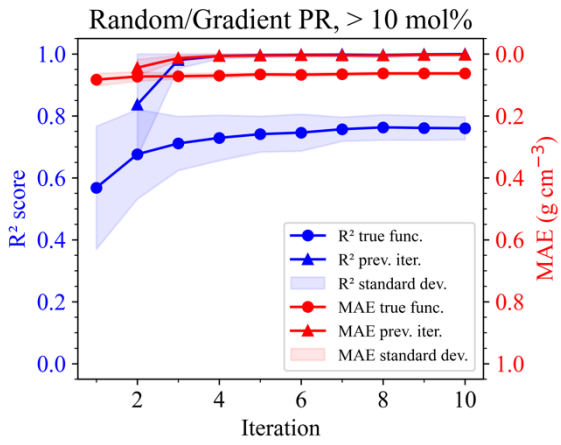

Figure S401: Evaluation of a compositional selection method (by the highest gradient of $\hat{\rho}_{\mathrm{SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using PR. A minimum distance criterion between compositions of 10 mol% is applied.

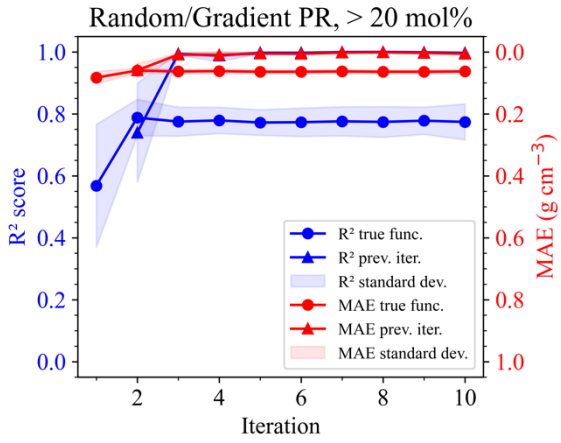

Figure S402: Evaluation of a compositional selection method (by the highest gradient of $\hat{\rho}_{\mathrm{SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using PR. A minimum distance criterion between compositions of 20 mol% is applied.

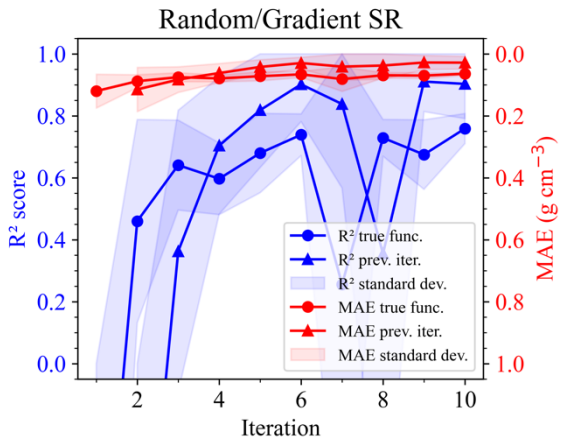

Figure S403: Evaluation of a compositional selection method (by the highest gradient of $\hat{\rho}_{\mathrm{SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using SR.

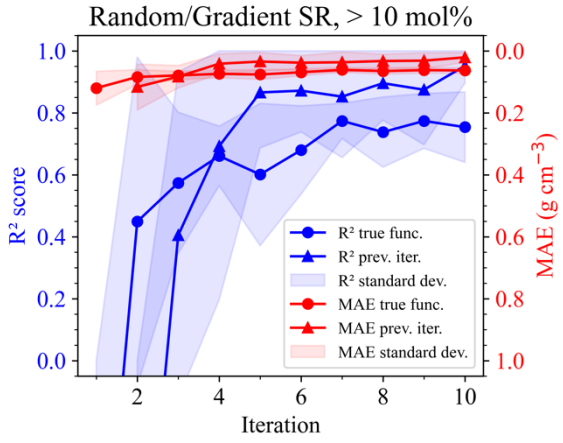

Figure S404: Evaluation of a compositional selection method (by the highest gradient of $\hat{\rho}_{\mathrm{SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using SR. A minimum distance criterion between compositions of 10 mol% is applied.

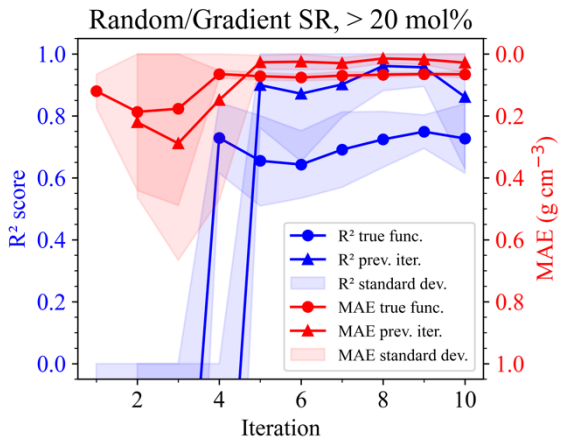

Figure S405: Evaluation of a compositional selection method (by the highest gradient of $\hat{\rho}_{\mathrm{SciGl}}(\mathbf{x})$) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using SR. A minimum distance criterion between compositions of 20 mol% is applied.

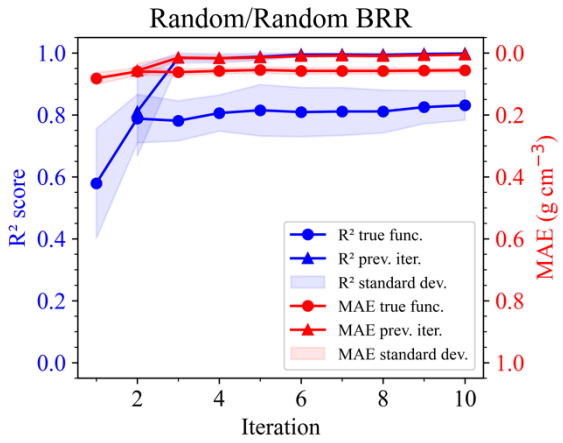

Figure S406: Evaluation of a compositional selection method (random selection) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using BRR.

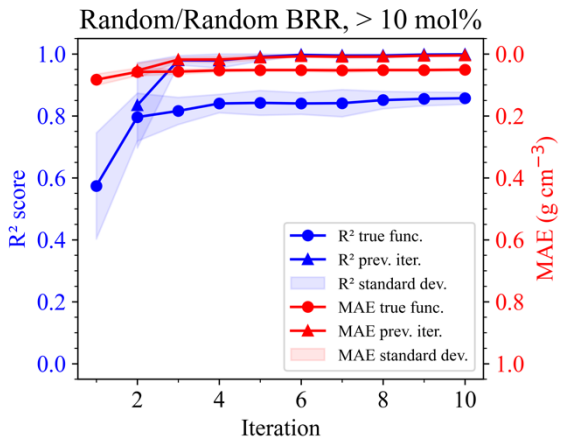

Figure S407: Evaluation of a compositional selection method (random selection) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using BRR. A minimum distance criterion between compositions of 10 mol% is applied.

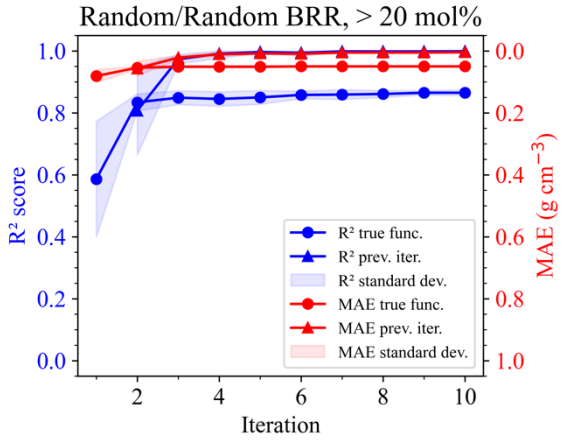

Figure S408: Evaluation of a compositional selection method (random selection) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using BRR. A minimum distance criterion between compositions of 20 mol% is applied.

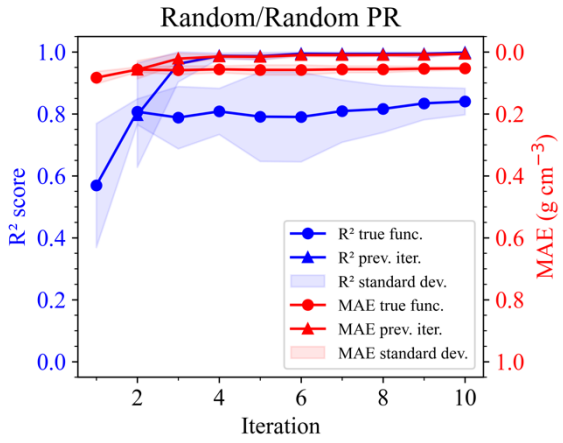

Figure S409: Evaluation of a compositional selection method (random selection) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using PR.

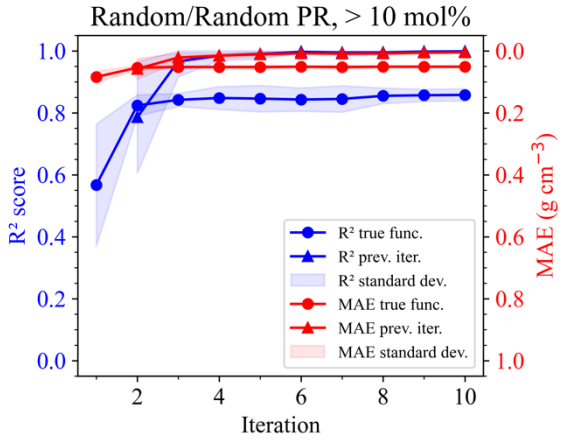

Figure S410: Evaluation of a compositional selection method (random selection) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using PR. A minimum distance criterion between compositions of 10 mol% is applied.

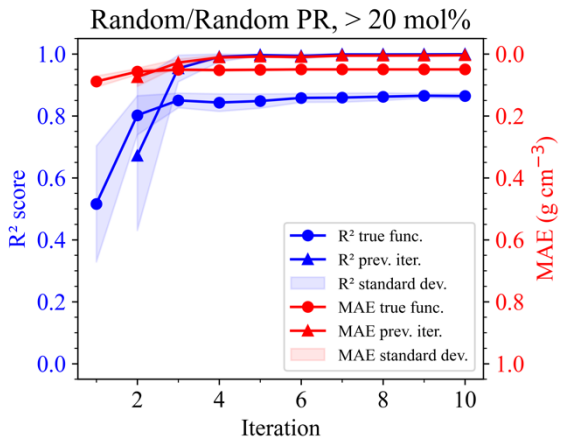

Figure S411: Evaluation of a compositional selection method (random selection) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using PR. A minimum distance criterion between compositions of 20 mol% is applied.

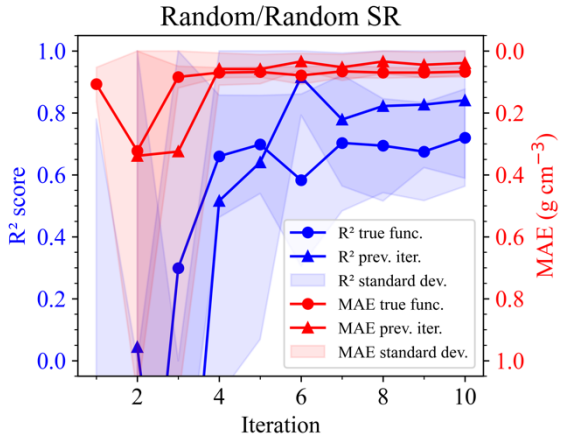

Figure S412: Evaluation of a compositional selection method (random selection) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using SR.

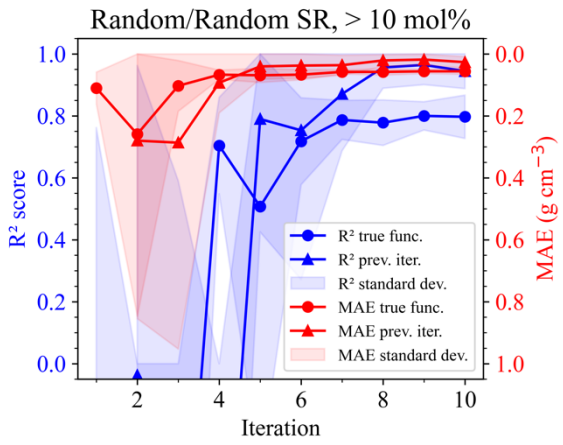

Figure S413: Evaluation of a compositional selection method (random selection) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using SR. A minimum distance criterion between compositions of 10 mol% is applied.

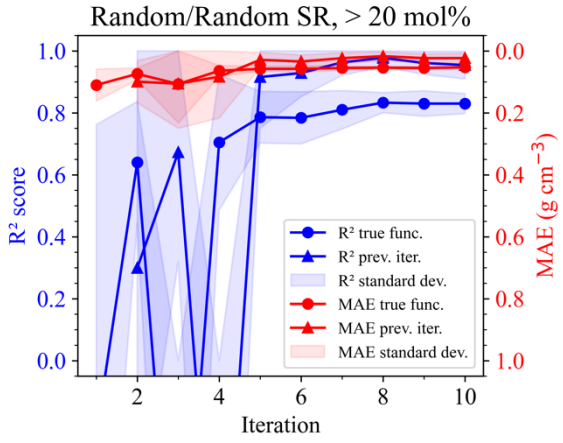

Figure S414: Evaluation of a compositional selection method (random selection) for iterative training of Simulation Calibrated Active Learning Estimator (SCALE) models, measured by R^2 score and MAE. The initial selection is random, the models are determined using SR. A minimum distance criterion between compositions of 20 mol% is applied.

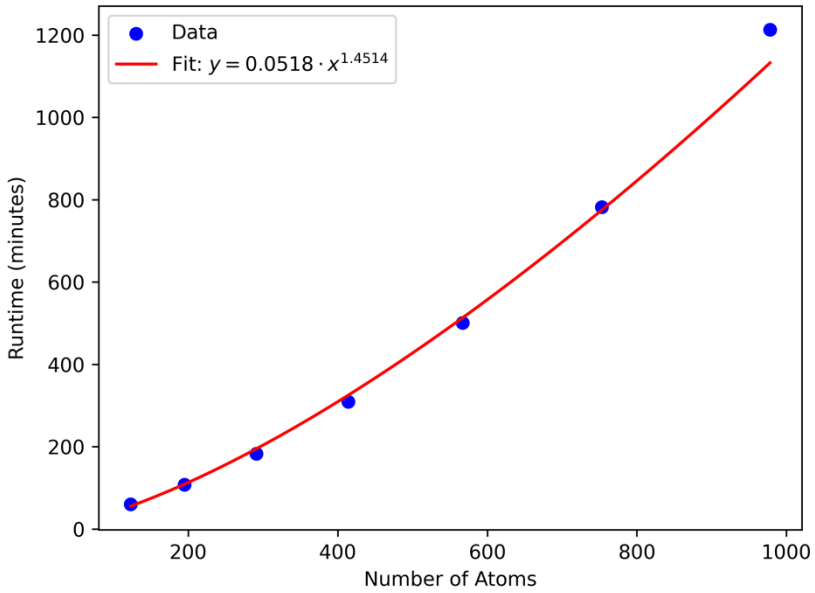

Figure S415: Runtime of MD (LAMMPS) simulations of SiO_2 with varying cell sizes and corresponding number of atoms. Each calculation is performed on a single Intel Xeon E5-2650 core with 2.2 GHz.

Table S23: Details of the GPR model simulating glass measurements from the robotic melting facility, trained on SciGlass data.

Implementation	Exact GPyTorch model, (Gardner, J.; Pleiss, G.; Weinberger, K. Q.; Bindel, D.; Wilson, A. G. Gpytorch: Blackbox Matrix-Matrix Gaussian Process Inference with Gpu Acceleration. *Adv. Neural Inf. Process. Syst.* **2018**, *31*.)
Prior	Zero-mean
Kernel function	Radial basis function (RBF)
Length scale	Smoothed box function (limits: 10^{-5} … 10^{5}), length scale prior 0.01
Noise prior	None
Learning rate	0.01
Max. training epochs	25
Dataset source	https://github.com/epam/SciGlass

Table S24: Selection of compositions melted and measured for Simulation Calibrated Active Learning Estimator (SCALE) iteration 1, including measured density, density obtained from MD simulations, and density predicted by the SCALE model using the density measurements from the first iteration.

Glass	Density measured (g cm^{-3})	Density simulation (g cm^{-3})	Density SCALE model 1 (g cm^{-3})
SiO_2	2.20	2.26	2.29
10 Al_2O_3 – 50 Na_2O – 40 SiO_2	2.57	2.75	2.51
50 Na_2O-50 SiO_2	2.56	2.68	2.61
10 Al_2O_3 – 15 B_2O_3 – 50 Na_2O – 25 SiO_2	2.42	2.78	2.57
10 Al_2O_3 – 20 B_2O_3 – 50 Na_2O – 20 SiO_2	2.92	2.77	2.56
7.5 Al_2O_3 – 50 Na_2O – 42.5 SiO_2	2.59	2.73	2.59

Table S25: Selection of compositions melted and measured for Simulation Calibrated Active Learning Estimator (SCALE) iteration 2, including measured density, density obtained from MD simulations, and density predicted by the SCALE model using the density measurements from iterations 1 and 2.

Glass	Density measured (g cm^{-3})	Density simulation (g cm^{-3})	Density SCALE model 2 (g cm^{-3})
10 Al_2O_3 – 40 B_2O_3 – 50 Na_2O	2.53	2.74	2.62
10 Al_2O_3 – 25 B_2O_3 – 50 Na_2O – 15 SiO_2	2.42	2.79	2.64

10 Al_2O_3 – 10 B_2O_3 – 50 Na_2O – 30 SiO_2	2.52	2.79	2.65
10 Al_2O_3 – 55 B_2O_3 – 35 Na_2O	2.36	2.72	2.63
10 Al_2O_3 – 40 B_2O_3 – 35 Na_2O – 15 SiO_2	2.47	2.71	2.63
10 Al_2O_3 – 42.5 Na_2O – 47.5 SiO_2	2.56	2.70	2.62

Table S26: Selection of compositions melted and measured for Simulation Calibrated Active Learning Estimator (SCALE) iteration 3, including measured density, density obtained from MD simulations, and density predicted by the SCALE model using the density measurements from iterations 1-3.

Glass	Density measured ($g\ cm^{-3}$)	Density simulation ($g\ cm^{-3}$)	Density SCALE model 3 ($g\ cm^{-3}$)
10 Al_2O_3 – 42.5 B_2O_3 – 47.5 Na_2O	2.47	2.72	2.52
10 Al_2O_3 – 57.5 B_2O_3 – 32.5 Na_2O	2.34	2.69	2.51
10 Al_2O_3 – 27.5 B_2O_3 – 47.5 Na_2O – 15 SiO_2	2.52	2.76	2.55
10 Al_2O_3 – 42.5 B_2O_3 – 32.5 Na_2O – 15 SiO_2	2.41	2.75	2.54

10 Al_2O_3 – 12.5 B_2O_3 – 47.5 Na_2O – 30 SiO_2	2.51	2.77	2.58
10 Al_2O_3 – 72.5 B_2O_3 – 17.5 Na_2O	2.14	2.62	2.42

Table S27: Selection of compositions yet to be melted and measured for Simulation Calibrated Active Learning Estimator (SCALE) iteration 4, including density obtained from MD simulations, and density predicted by the SCALE model using the density measurements from iterations 1-3.

Glass	Density simulation (g cm^{-3})	Density SCALE model 4 (g cm^{-3})
10 Al_2O_3 – 70 B_2O_3 – 20 Na_2O	2.65	2.29
10 Al_2O_3 – 55 B_2O_3 – 35 Na_2O	2.72	2.39
10 Al_2O_3 – 85 B_2O_3 – 5 Na_2O	2.27	1.94
10 Al_2O_3 – 55 B_2O_3 – 20 Na_2O – 15 SiO_2	2.66	2.37
10 Al_2O_3 – 37.5 B_2O_3 – 50 Na_2O – 2.5 SiO_2	2.74	2.45
10 Al_2O_3 – 70 B_2O_3 – 5 Na_2O – 15 SiO_2	2.28	2.01

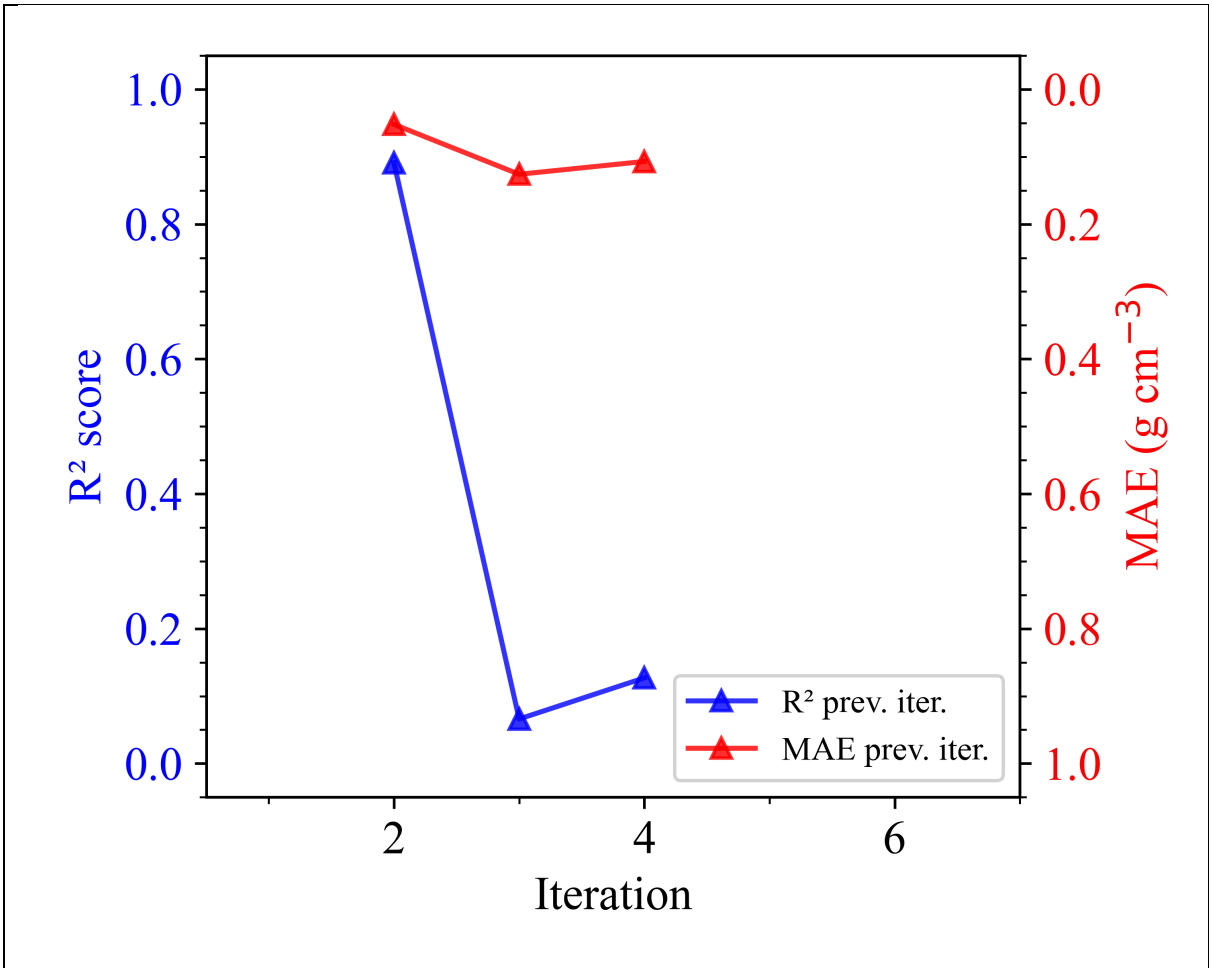

Figure S416: Comparison of the similarity between consecutive Simulation Calibrated Active Learning Estimator (SCALE) models to evaluate convergence.

Table S28: Measured densities of NABS glass.

Glass	Density measured ($g\ cm^{-3}$)
25 B_2O_3 – 25 Na_2O – 50 SiO_2	2.5345
35 B_2O_3 – 15 Na_2O – 50 SiO_2	2.3597

40 B_2O_3 – 10 Na_2O – 50 SiO_2	2.2047
30 B_2O_3 – 20 Na_2O – 50 SiO_2	2.4752
10 B_2O_3 – 16 Na_2O – 74 SiO_2	2.4712
75 B_2O_3 – 10 Na_2O – 15 SiO_2	2.1670
73 B_2O_3 – 7 Na_2O – 20 SiO_2	2.0549
68 B_2O_3 – 7 Na_2O – 25 SiO_2	2.0480
63 B_2O_3 – 7 Na_2O – 30 SiO_2	2.0556
58 B_2O_3 – 7 Na_2O – 35 SiO_2	2.0745
53 B_2O_3 – 7 Na_2O – 40 SiO_2	2.0919
48 B_2O_3 – 7 Na_2O – 45 SiO_2	2.1030
43 B_2O_3 – 7 Na_2O – 50 SiO_2	2.1255
38 B_2O_3 – 7 Na_2O – 55 SiO_2	2.1440

33 B_2O_3 – 7 Na_2O – 60 SiO_2	2.1642
28 B_2O_3 – 7 Na_2O – 65 SiO_2	2.1903
23 B_2O_3 – 7 Na_2O – 70 SiO_2	2.2092
18 B_2O_3 – 7 Na_2O – 75 SiO_2	2.2535
6 Al_2O_3 – 35 B_2O_3 – 35 Na_2O – 24 SiO_2	2.4526
3 Al_2O_3 – 61 B_2O_3 – 11 Na_2O – 25 SiO_2	2.1006
6 Al_2O_3 – 60 B_2O_3 – 10 Na_2O – 24 SiO_2	2.1114
9 Al_2O_3 – 59 B_2O_3 – 9 Na_2O – 23 SiO_2	2.1287
1.5 Al_2O_3 – 61.5 B_2O_3 – 11.5 Na_2O – 25.5 SiO_2	2.1233
7.5 Al_2O_3 – 59.5 B_2O_3 – 9.5 Na_2O – 23.5 SiO_2	2.1183
4.5 Al_2O_3 – 60.5 B_2O_3 – 10.5 Na_2O – 24.5 SiO_2	2.1090
22.5 Al_2O_3 –22.5 Na_2O – 55 SiO_2	2.4772

30 Al_2O_3 – 5 B_2O_3 – 20 Na_2O – 45 SiO_2	2.4651
35 Al_2O_3 – 5 B_2O_3 – 20 Na_2O – 40 SiO_2	2.5133
32.5 Al_2O_3 – 5 B_2O_3 – 20 Na_2O – 42.5 SiO_2	2.4822
4 Al_2O_3 – 4.2 B_2O_3 – 13.6 Na_2O – 78.2 SiO_2	2.3831
6 Al_2O_3 – 13 B_2O_3 – 13 Na_2O – 68 SiO_2	2.3826
6 Al_2O_3 – 15.5 B_2O_3 – 15.5 Na_2O – 63 SiO_2	2.3906
3 Al_2O_3 – 14 B_2O_3 – 14 Na_2O – 69 SiO_2	2.4228
17.5 B_2O_3 – 17.5 Na_2O – 65 SiO_2	2.4928
20 B_2O_3 – 20 Na_2O – 60 SiO_2	2.5060
72 B_2O_3 – 2 Na_2O – 26 SiO_2	1.9294
69.5 B_2O_3 – 4.5 Na_2O – 26 SiO_2	1.9936
67 B_2O_3 – 7 Na_2O – 26 SiO_2	2.0436

64.5 B_2O_3 – 9.5 Na_2O – 26 SiO_2	2.0979
62 B_2O_3 – 12 Na_2O – 26 SiO_2	2.1528
59.5 B_2O_3 – 14.5 Na_2O – 26 SiO_2	2.2057
57 B_2O_3 – 17 Na_2O – 26 SiO_2	2.2507
54.5 B_2O_3 – 19.5 Na_2O – 26 SiO_2	2.3102
52 B_2O_3 – 22 Na_2O – 26 SiO_2	2.3527
49.5 B_2O_3 – 24.5 Na_2O – 26 SiO_2	2.4023
47 B_2O_3 – 27 Na_2O – 26 SiO_2	2.4356
44.5 B_2O_3 – 29.5 Na_2O – 26 SiO_2	2.4631
39.5 B_2O_3 – 34.5 Na_2O – 26 SiO_2	2.4781
34.5 B_2O_3 – 39.5 Na_2O – 26 SiO_2	2.4670
32 B_2O_3 – 42 Na_2O – 26 SiO_2	2.4760

29.5 B_2O_3 – 44.5 Na_2O – 26 SiO_2	2.4733
15 B_2O_3 – 15 Na_2O – 70 SiO_2	2.4350
33.3 B_2O_3 – 33.3 Na_2O – 33.4 SiO_2	2.5364
37 B_2O_3 – 37 Na_2O – 26 SiO_2	2.4793
13 B_2O_3 – 13 Na_2O – 74 SiO_2	2.4448

Table S29: Atomic charges, Morse potential parameters (D_{ij}, α_{ij}, r^o_{ij}), and repulsive parameters (B_{ij}) for pairwise interactions among constituents of NABS (Na_2O–Al_2O_3–B_2O_3–SiO_2) glass, as defined by the BMP (Bertani–Menziani–Pedone; Bertani, M.; Pallini, A.; Cocchi, M.; Menziani, M. C.; Pedone, A. A New Self-Consistent Empirical Potential Model for Multicomponent Borate and Borosilicate Glasses. *J. Am. Ceram. Soc.* **2022**, *105* (12), 7254–7271. https://doi.org/10.1111/jace.18681) potential. These parameters are used in atomistic MD simulations to model interatomic forces in multicomponent oxide glasses.

Atom pair	$\mathbf{D_{ij}}$ **(eV)**	$\mathbf{a_{ij}}$ **(Å^{-2})**	$\mathbf{r_{ij}^0}$ **(Å)**	$\mathbf{B_{ij}}$ **(r^{12})**
$Si^{2.4}$ - $O^{-1.2}$	0.3406	2.0067	2.1000	1.0
$Na^{0.6}$ - $O^{-1.2}$	0.0234	1.7639	3.0063	5.0
$Al^{1.8}$ - $O^{-1.2}$	0.3616	1.9004	2.1648	0.9
$B^{1.8}$ - $O^{-1.2}$	Eq. S2	2.6433	1.4367	10.0
$O^{-1.2}$ - $O^{-1.2}$	0.0424	1.3793	3.6187	100.0

Table S30: Buckingham potential parameters (A_{ij}, ρ_{ij}) for pairwise interactions among constituents of NABS ($Na_2O–Al_2O_3–B_2O_3–SiO_2$) glass, as defined by the BMP (Bertani–Menziani–Pedone; Bertani, M.; Pallini, A.; Cocchi, M.; Menziani, M. C.; Pedone, A. A New Self-Consistent Empirical Potential Model for Multicomponent Borate and Borosilicate Glasses. *J. Am. Ceram. Soc.* **2022**, *105* (12), 7254–7271. https://doi.org/10.1111/jace.18681) potential. These parameters are used in atomistic Molecular Dynamics simulations to model interatomic forces in multicomponent oxide glasses.

Atom pair	$\mathbf{A_{ij}}$ **(eV)**	$\boldsymbol{\rho_{ij}}$ **(Å^{-2})**
Si-Si	7.0937	0.9756
Si-B	8.9594	0.9270
B-B	8.9594	0.8012

Table S31: Three-body-potential parameters (K_{ijk}, θ^0_{ijk}, ρ_{tb}) for interactions among constituents of NABS ($Na_2O–Al_2O_3–B_2O_3–SiO_2$) glass, as defined by the BMP (Bertani–Menziani–Pedone; Bertani, M.; Pallini, A.; Cocchi, M.; Menziani, M. C.; Pedone, A. A New Self-Consistent Empirical Potential Model for Multicomponent Borate and Borosilicate Glasses. *J. Am. Ceram. Soc.* **2022**, *105* (12), 7254–7271. https://doi.org/10.1111/jace.18681) potential. These parameters are used in atomistic Molecular Dynamics simulations to model interatomic forces in multicomponent oxide glasses.

Atom triplet	$\mathbf{K_{ijk}}$ **(eV rad^{-2})**	$\boldsymbol{\theta^0_{ijk}}$ (deg)	$\boldsymbol{\rho_{tb}}$ (Å)
Si-O-Si	25.0	109.47	1
B-O-B	60.0	109.47	1

B-O-Si	60.0	109.47	1

Force field equations

PMMCS (Pedone, A.; Malavasi, G.; Menziani, M. C.; Cormack, A. N.; Segre, U. A New Self-Consistent Empirical Interatomic Potential Model for Oxides, Silicates, and Silica-Based Glasses. *J. Phys. Chem. B* **2006**, *110* (24), 11780–11795. https://doi.org/10.1021/jp0611018.)

BMP (Bertani–Menziani–Pedone; Bertani, M.; Pallini, A.; Cocchi, M.; Menziani, M. C.; Pedone, A. A New Self-Consistent Empirical Potential Model for Multicomponent Borate and Borosilicate Glasses. *J. Am. Ceram. Soc.* **2022**, *105* (12), 7254–7271. https://doi.org/10.1111/jace.18681.)

The PMMCS force field formula can be expressed as follows:

$$\mathrm{U}(r_{ij}) = \frac{\mathrm{z}_i \mathrm{z}_j \mathrm{e}^2}{r_{ij}} + \mathrm{D}_{ij}\left[\left(1 - \mathrm{e}^{-\mathrm{a}_{ij}\left(\mathrm{r}_{ij} - \mathrm{r}_{ij}^0\right)}\right)^2 - 1\right] + \frac{\mathrm{B}_{ij}}{r_{ij}^{12}} \tag{S1}$$

with D_{ij}, a_{ij}, and r_{ij}^0 being the Morse function coefficients and z_i and z_j corresponding to the partial charges of ions i and j respectively. The Morse parameter D_{ij} for the B-O interaction is calculated as follows:

$$\mathrm{D}_{\mathrm{B-O}} = \mathrm{a}(R) \cdot \mathrm{e}^{\mathrm{b}(R,K)} + \mathrm{c}(R) \cdot \mathrm{e}^{\mathrm{d}(R,K)} + 0.001665 \cdot K^3 - 0.12807 \cdot R$$

$$\mathrm{a}(R) = 2.11081 + \frac{1}{(R+1)^2}$$

$$\mathrm{b}(R,K) = 0.02063 \cdot R + 0.06312 \cdot K \tag{S2}$$

$$\mathrm{c}(R) = -2.12213 + \frac{1}{(R+1)^2}$$

$$\mathrm{d}(R,K) = -7.50152 \cdot R + 0.32778 \cdot K$$

The equations for Buckingham potentials and three-body-screened-harmonic (shrm) interactions are as follows:

$$\mathrm{U}(r_{ij}) = \mathrm{A}_{ij}\mathrm{e}^{-\frac{r_{ij}}{\rho_{ij}}} \tag{S3}$$

$$\mathrm{U}(\theta_{ijk}) = \frac{\mathrm{K}_{ijk}}{2}(\theta_{ijk} - \theta_{ijk}^0)^2 \mathrm{e}^{-\left(\frac{\mathrm{r}_{ij}}{\rho_{\mathrm{tb}}} + \frac{\mathrm{r}_{jk}}{\rho_{\mathrm{tb}}}\right)} \tag{S4}$$

where A_{ij} and ρ_{ij} are Buckingham potential parameters while K_{ijk}, $\theta_{ijk,0}$ and ρ_{tb} are three-body parameters.

Table S32: Details of HT-MD simulations for NABS glasses. For each composition, properties were computed from 10 independently generated cells. Cells were regenerated and re-simulated if a property estimate lay outside the Tukey fence [Q1−2×IQR,Q3+2×IQR] [Q1−2×IQR, Q3+2×IQR]. Here, Q1 and Q3 denote the first and third quartiles and IQR=Q3−Q1IQR=Q3−Q1. V14 and V18 refer to the two simulated cell sizes.

Stage	**Simulated cells, V14**	**Simulated cells, V18**
Initial	4370	28520
Re-run 1	1233	7393
Re-run 2	171	885

Table S33: Final dataset size for the V14 and V18 HT-MD datasets. A composition was retained in the final dataset only if at least 70% of its computed property values fell within [Q1−2×IQR,Q3+2×IQR] [Q1−2×IQR, Q3+2×IQR] (with Q1, Q3, and IQR defined as in Table S32).

	V14	**V18**
Possible compositions	437	2852
Compositions retained	410	2549