

## Appendix A. Additional dataset results

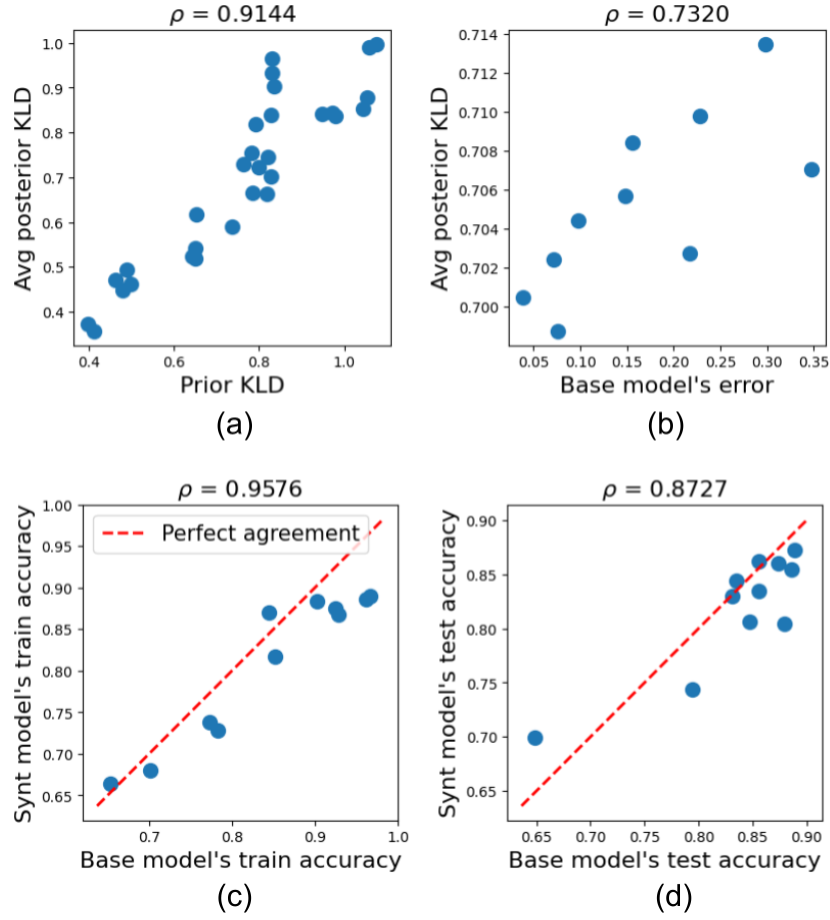


Figure A.1: Correlation plots and coefficients illustrating key relationships in our inverse estimation analysis for a **deep neural network model** trained on the **Retinal Ganglion Cell Stimulus Types** dataset.

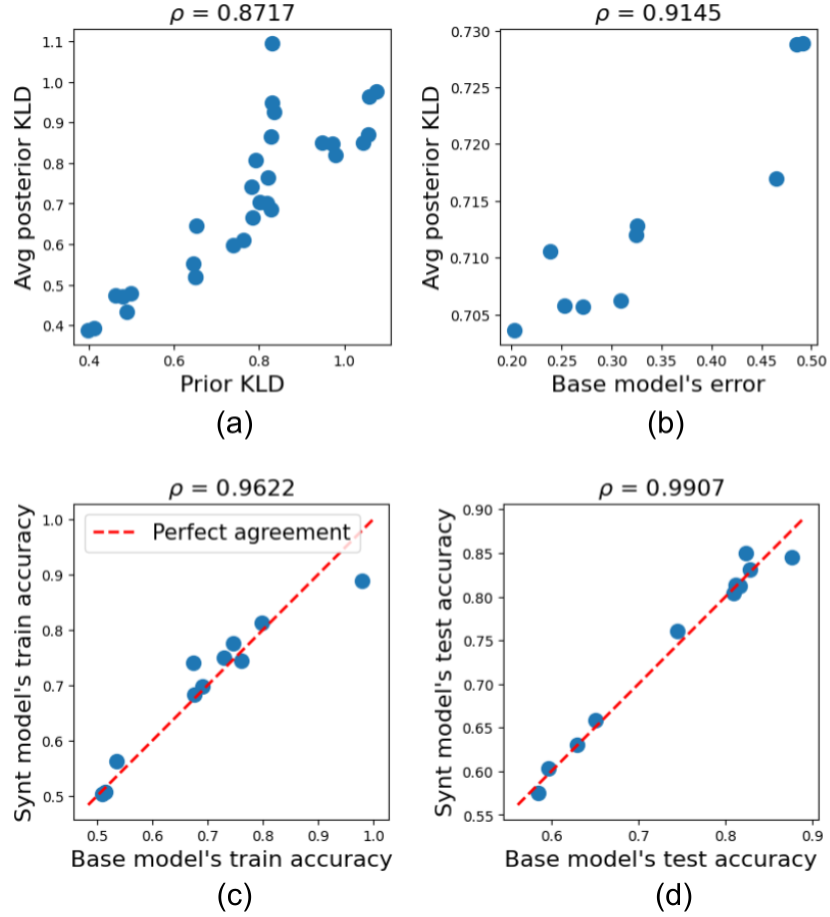


Figure A.2: Correlation plots and coefficients illustrating key relationships in our inverse estimation analysis for a **gradient boosting machine model** trained on the **Retinal Ganglion Cell Stimulus Types** dataset.

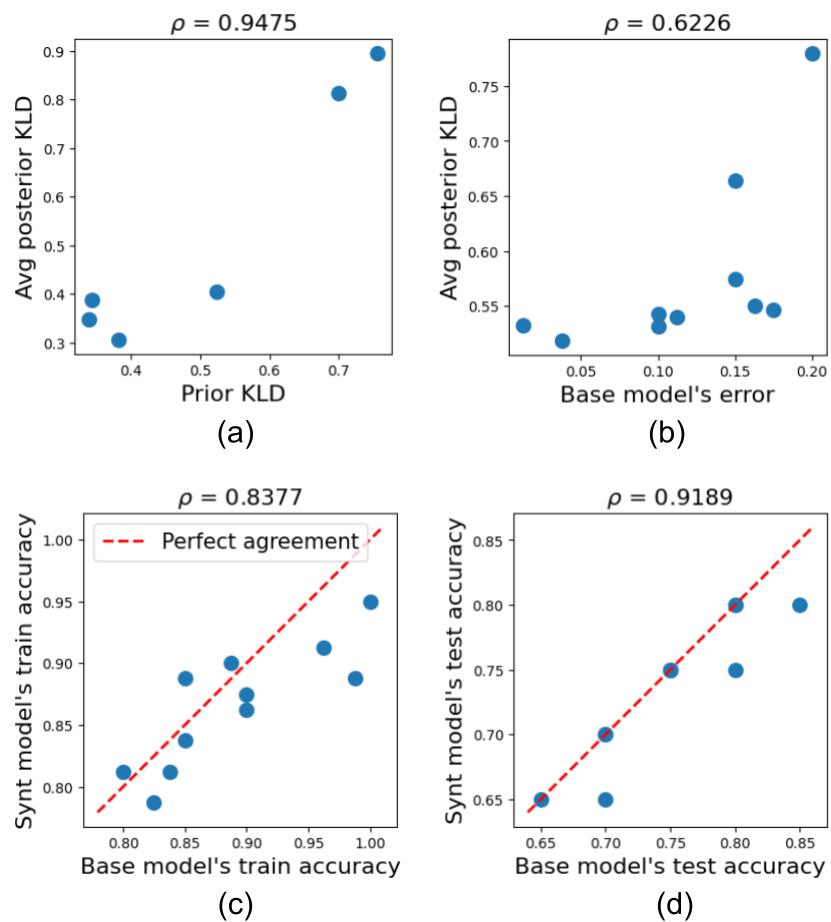


Figure A.3: Correlation plots and coefficients illustrating key relationships in our inverse estimation analysis for a **deep neural network model** trained on the **Sirtuin6 Small Molecules dataset**.

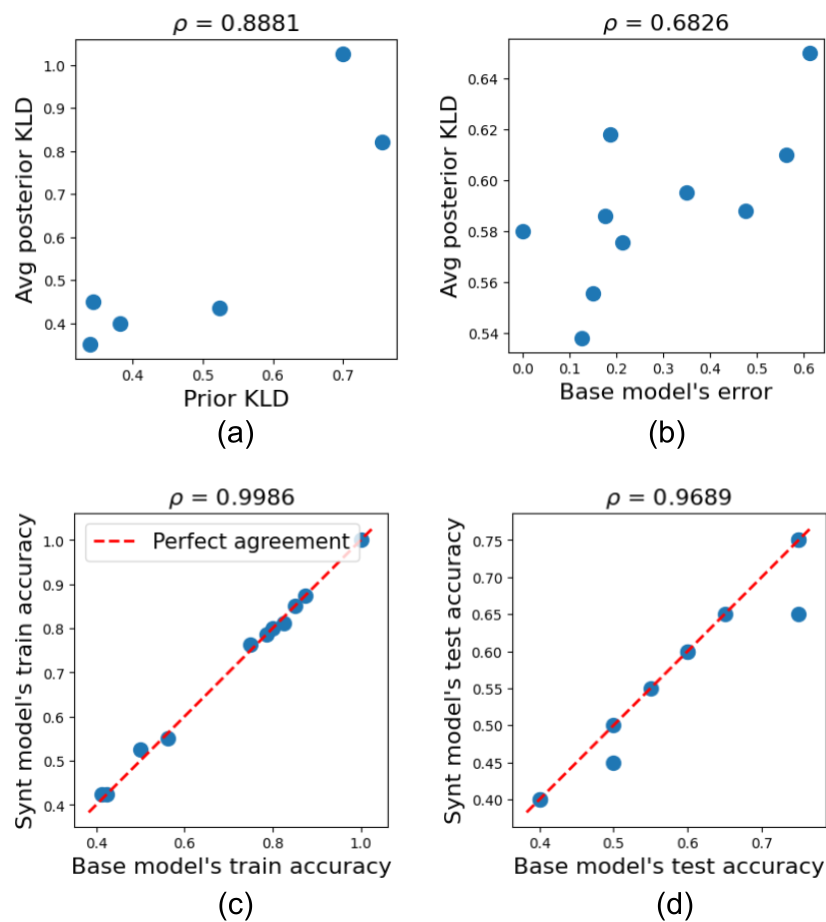


Figure A.4: Correlation plots and coefficients illustrating key relationships in our inverse estimation analysis for a **decision tree classifier model** trained on the **Sirtuin6 Small Molecules dataset**.

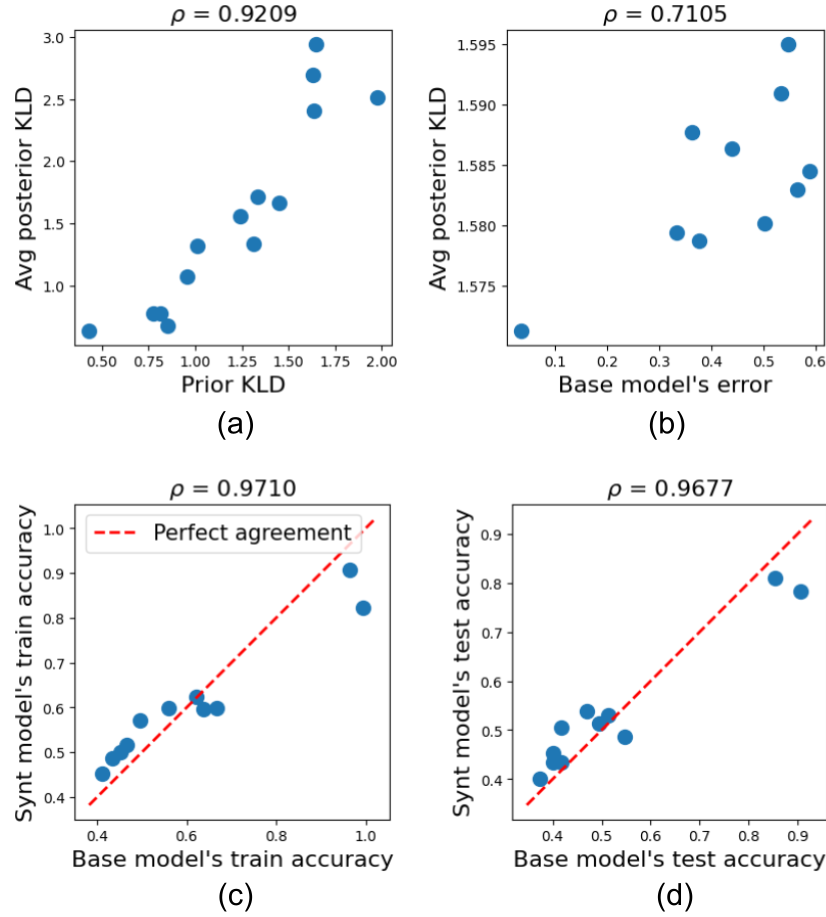


Figure A.5: Correlation plots and coefficients illustrating key relationships in our inverse estimation analysis for a **deep neural network model** trained on the **Anemia Types dataset**.

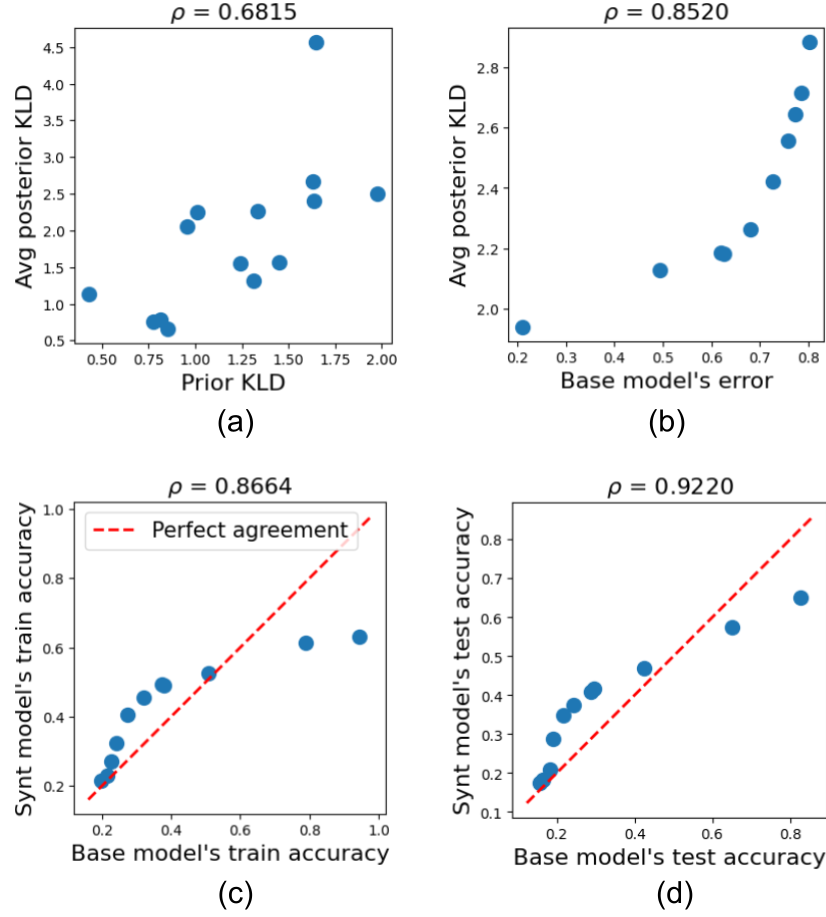


Figure A.6: Correlation plots and coefficients illustrating key relationships in our inverse estimation analysis for a **voting classifier model** trained on the **Anemia Types dataset**.

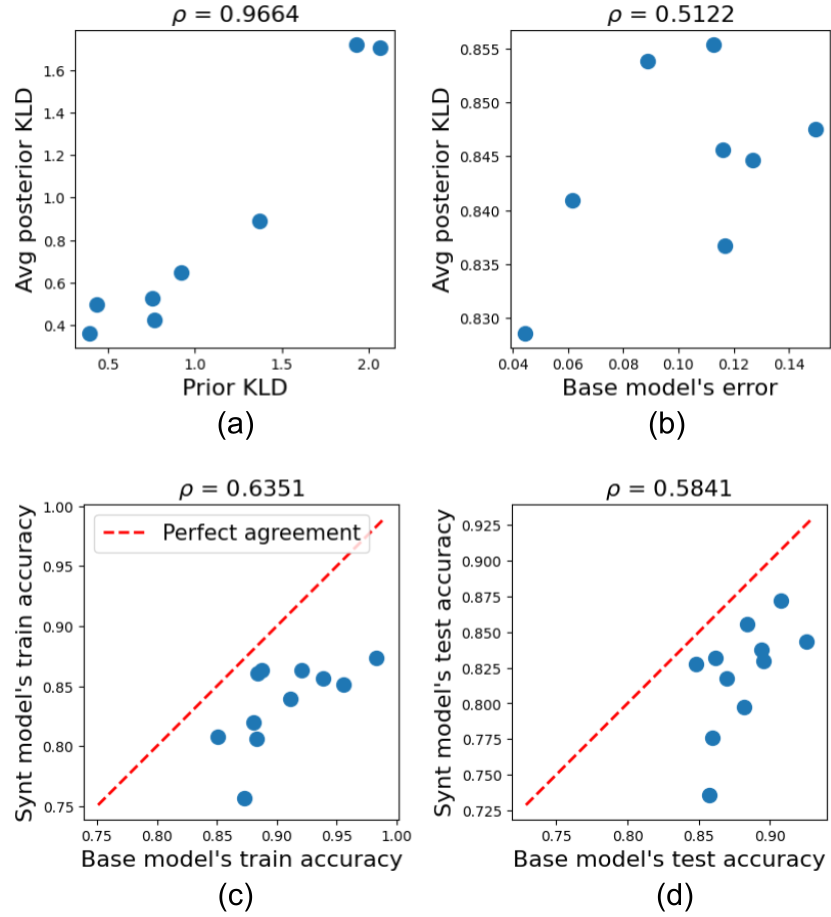


Figure A.7: Correlation plots and coefficients illustrating key relationships in our inverse estimation analysis for a **deep neural network model** trained on the **Brain Stroke dataset**.

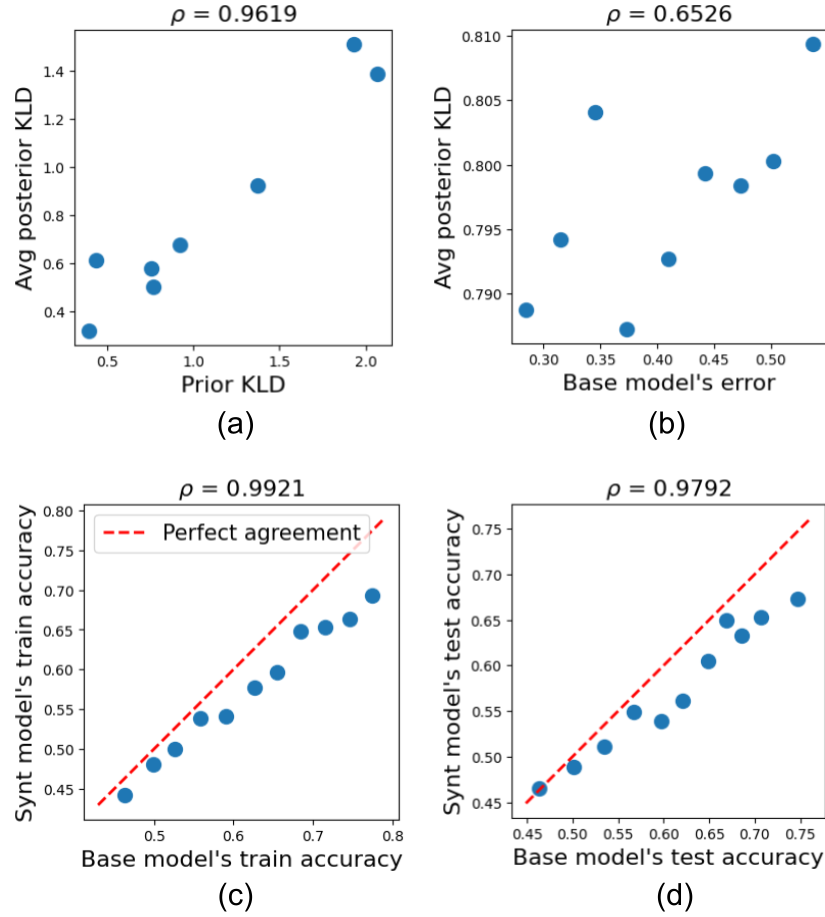


Figure A.8: Correlation plots and coefficients illustrating key relationships in our inverse estimation analysis for an **SVM** model trained on the **Brain Stroke** dataset.



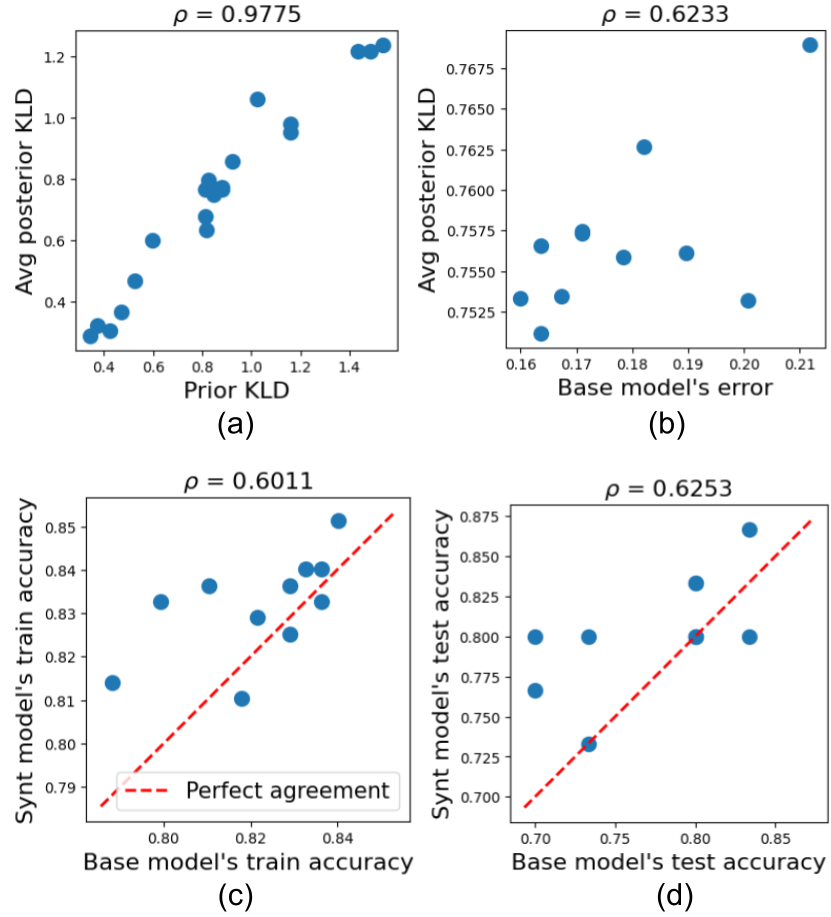


Figure A.9: Correlation plots and coefficients illustrating key relationships in our inverse estimation analysis for a **naive Bayes** model trained on the **Heart Disease** dataset.

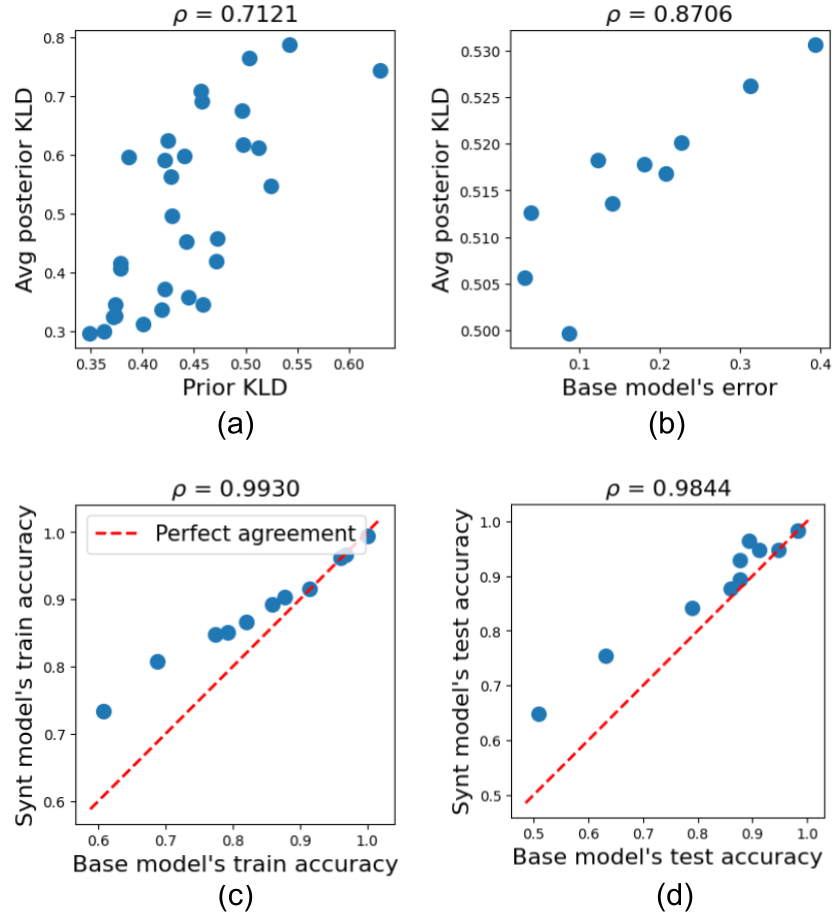


Figure A.10: Correlation plots and coefficients illustrating key relationships in our inverse estimation analysis for a **deep neural network model** trained on the **WBCD dataset**.

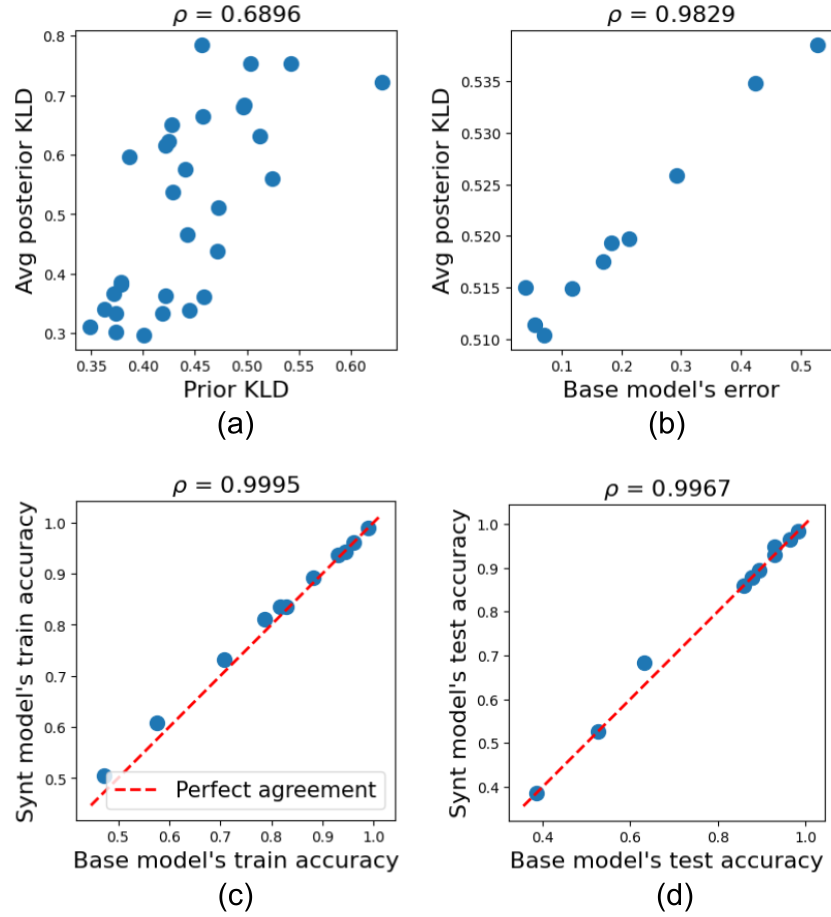


Figure A.11: Correlation plots and coefficients illustrating key relationships in our inverse estimation analysis for a **logistic regression model** trained on the **WBCD dataset**.

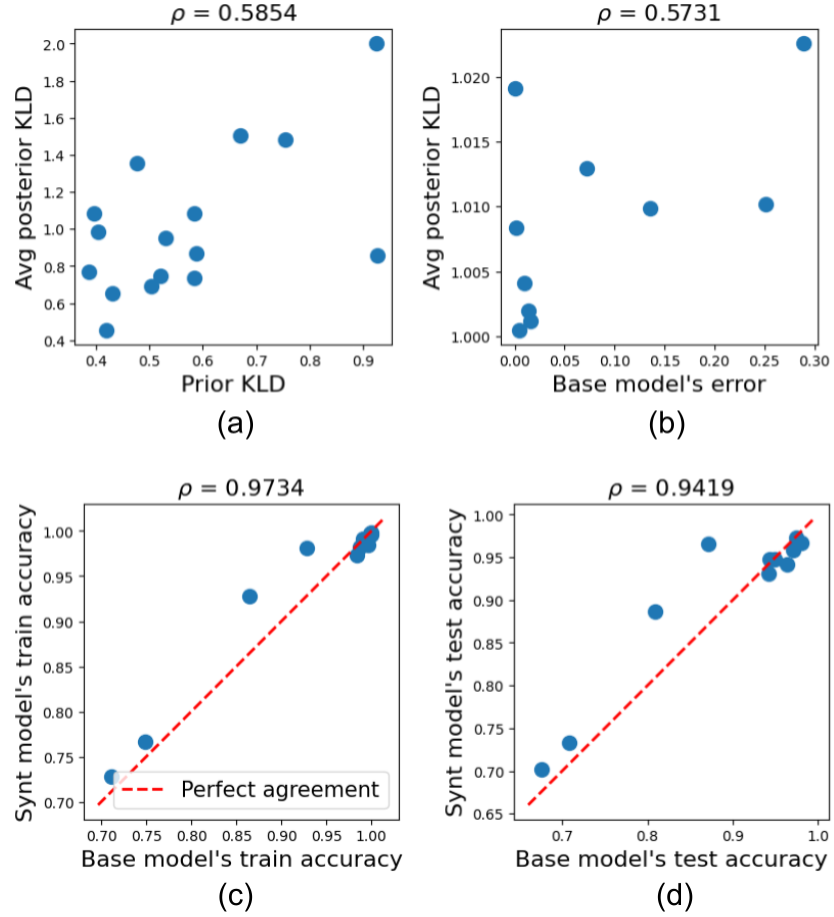


Figure A.12: Correlation plots and coefficients illustrating key relationships in our inverse estimation analysis for a **deep neural network model** trained on the **PenDigits (3 classes) dataset**.

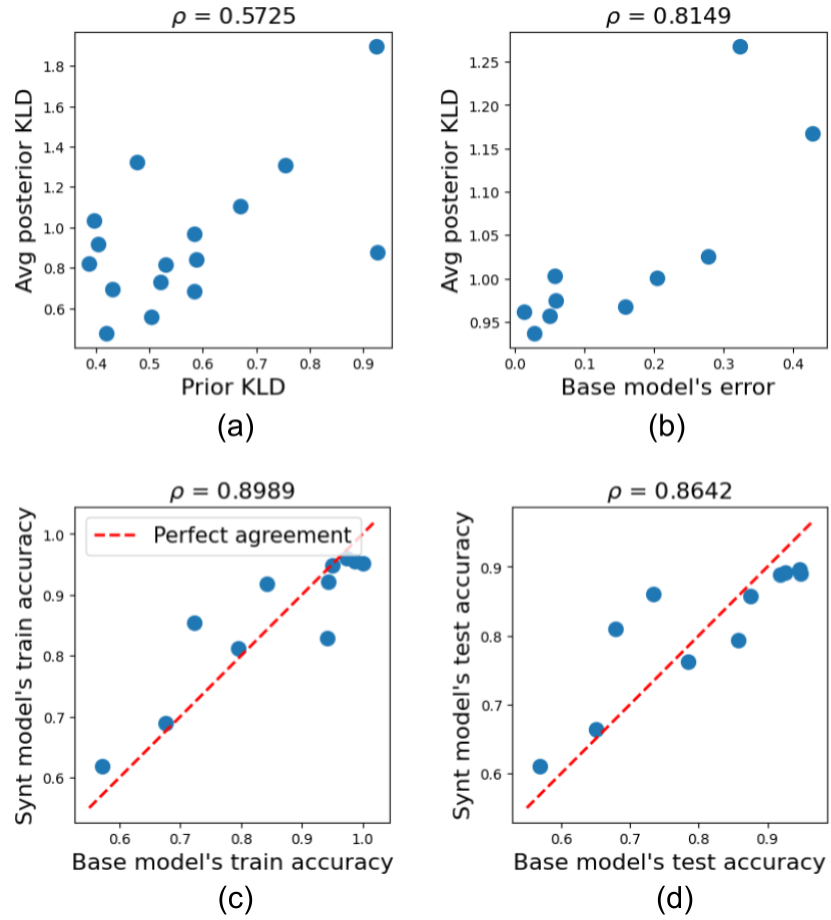


Figure A.13: Correlation plots and coefficients illustrating key relationships in our inverse estimation analysis for a **random forest model** trained on the **PenDigits (3 classes) dataset**.

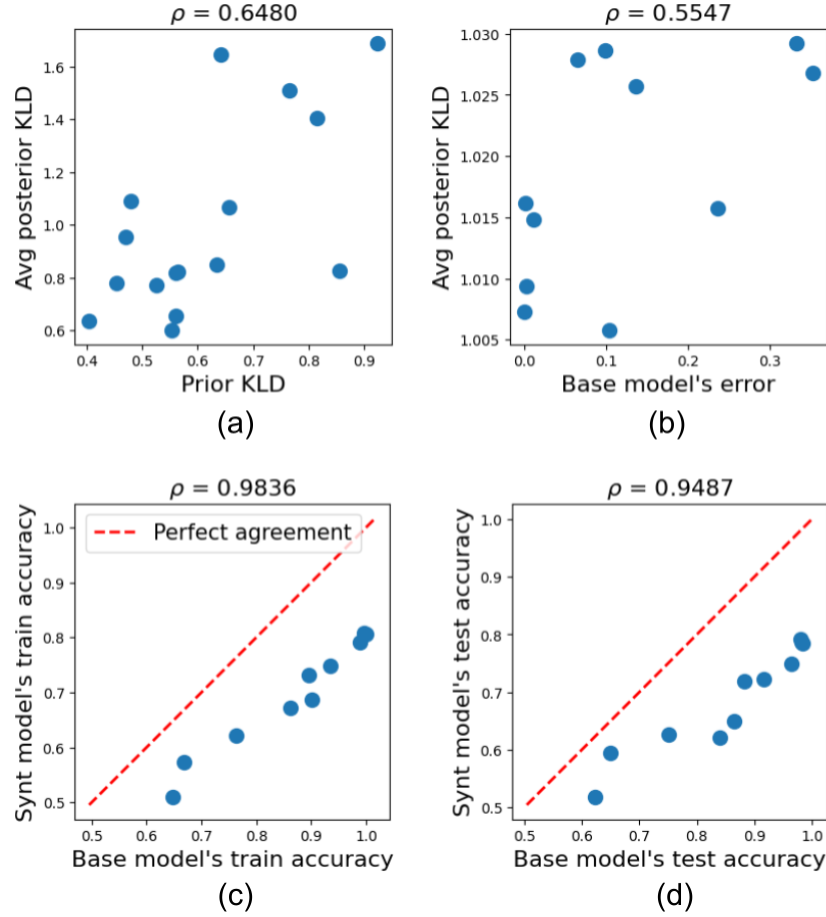


Figure A.14: Correlation plots and coefficients illustrating key relationships in our inverse estimation analysis for a **deep neural network model** trained on the **PenDigits (5 classes) dataset**.

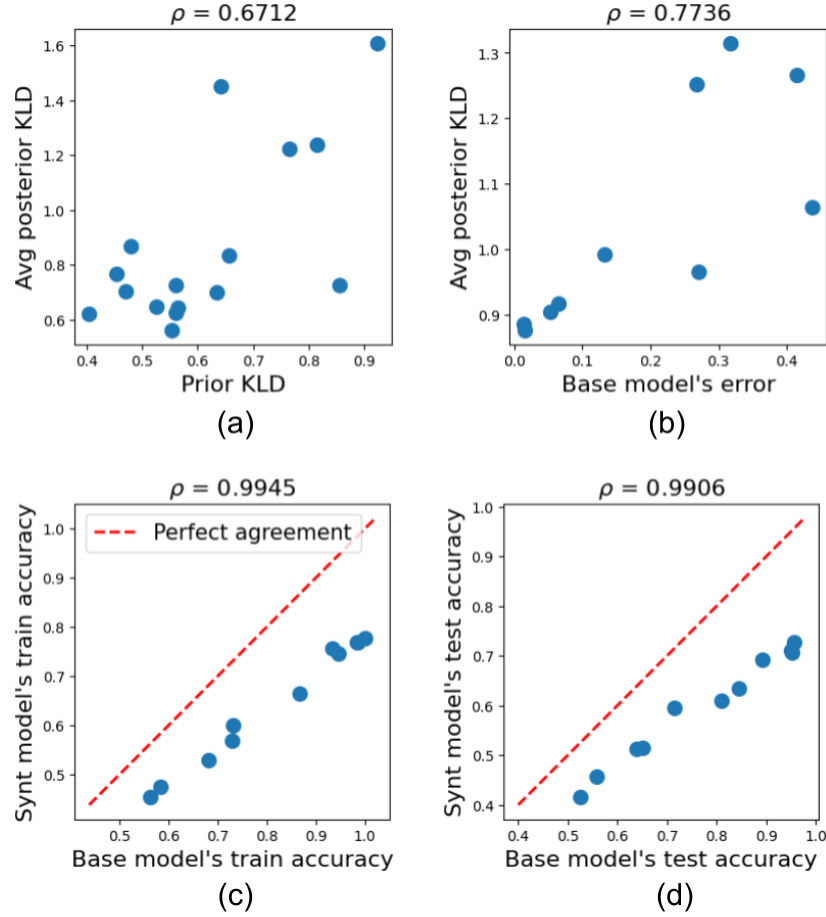


Figure A.15: Correlation plots and coefficients illustrating key relationships in our inverse estimation analysis for a **random forest model** trained on the **PenDigits (5 classes) dataset**.