

Code System

Code System	Memo	Frequency
Code System		293
Explainability purposes	Why do participants use ML explainability techniques like PDP and ICE plots? What are they interested in?	0
Guiding feature engineering		1
Understanding feature-outcome relationships	Trends between features and the model's output. Are the feature-outcome relationships intuitive?	11
Understanding feature importance		2
Identifying interactions		2
Sensitivity analysis		2
Model monitoring		1
Understanding null values		2
Explainability techniques	The participants' current techniques for ML explainability.	0
Convention and familiarity		4
Describing differences between clusters		2
Residual analysis	Performing residual analysis, such as using surrogate decision trees.	4
SHAP		6
ALE		2
Marginal plots		2
Pain point in comparing and understanding different types of plots		1
Uses scatter plots to identify interactions		3
Don't go in depth on interactions		5
Comparing feature vs. label plots for training and validation		1
PDP/ICE	These codes cover participants' current usage of PDP and ICE plots.	0
Don't primarily look at two-way plots		6
Uses ICE plots, but not PDPs		1
Does not use PDPs or ICE plots		2
Uses PDPs and ICE plots		2
Uses PDPs but not ICE plots		3
Which features to look at	These codes are about how the participants determine which features to examine in their current ML explainability techniques.	0
Already have features in mind		2
Earth mover's distance		1
Most important features		6
Comparing different types of plots for the same feature		1
Explores correlated features		1
No good approach to determining which plots to look at		1
PDPilot General Usage	This collection of codes is for general information about participants' experience using PDPilot, without being specific to a particular technique or part of the tool.	0
Over-reliance on guidance		1
Couldn't find use for technique	The participant did not find a situation to use a particular technique, even if they expected that technique to be useful	4
Didn't know how to do something in PDPilot	The participant did not know how to do something in PDPilot or forgot or misunderstood how a technique works.	7
Tool complexity		11
PDPilot Techniques Usage	These codes are about how participants used PDPilot, with an emphasis on the ranking, filtering, and clustering.	0
Interactions	These codes are for how PDPilot supports analyzing interactions, apart from the technique specific codes above.	0
Not very useful		1
Good for analyzing interactions		3
Liked interaction plots		2
Ranking	These codes are specific to the different ranking metrics.	0
Liked importance ranking		7
Liked cluster difference		1
Didn't use cluster difference		3
Highlighting was useful	This code applies to comments on the usefulness of brushing ICE lines.	10
Liked highlighted line similarity		3
Didn't use highlighted line similarity		5
Liked highlighted histogram difference		4
Didn't use highlighted histogram difference		2
Did not find sorting by interactions to be useful		1
Filtering	These codes apply to filtering, though participants did not talk much about filtering by shape.	0
Filtering was useful	Comments about the usefulness of filtering, though not specific to filtering by shape.	5
Filtering by shape		1
Liked two-way filters		3
Didn't use one-way filters		1
Clustering	These codes apply to clustering and comparisons between using the cluster descriptions vs. brushing and the ranking metrics.	0
Clusters and cluster descriptions were useful		4
Prefered highlighting over clusters		3
Prefered clusters over highlighting		1
PDPilot improvements	This group of codes is for suggestions and ideas for how to improve PDPilot.	0
Outliers, overfitting, and skewed distributions	These codes deal with concerns and strategies about dealing with outliers, skewed feature distributions, and the model overfitting.	0
Problem with few instances		3
Outliers affecting rankings		4
Removing outliers		3
Excluding features		1
Iterative approach to handling overfitting		2
Weighting metrics by data distribution		3
Concerned about correlated variables		3
Are plots trustworthy		1
Alternative to two-way scatter plot		1
Using different importance metric		3
Visualization sizing issue		5
Saving and exporting findings		2
Confused by cluster description stacked bar chart		2
Easier to highlight clusters		2
Incorporating information about errors		6
Handling null values		1
Study limitations	These codes are for issues that reflect limitations in the user study design.	0
Uninteresting dataset		3
Unfamiliar dataset		3
Didn't have a goal in mind		4
Benefits of PDPilot	Benefits of PDPilot and general compliments.	0
Exploratory vs. hypothesis driven approach		6
Pre-computed		1
Compliment		17
Creating plots by hand		2
Better at identifying heterogenous trends		2
Interaction	Comments on feature interaction	21
Anonymize	Code to anonymize segments.	32