HOMEWORK-2 IS733-DATAMINING

Siddhidhatri Rohith Reddy J(LO55859)

Part 1. Reflections on Homework 1

From the feedback you received (either from the instructor/TA or peers), what are the takeaways/lessons learned you could apply to future analysis?

Based on the feedback received, there are several valuable lessons to apply to future analysis and presentations:

1. **Excluding Code**: The feedback emphasizes the importance of excluding raw code from presentations. This lesson highlights the need for clear, concise communication and the use of visual aids to convey the key findings and insights rather than overwhelming the audience with technical details.

2. **Consider Alternative Visuals**: Suggestions about using tables or different types of plots indicate the importance of choosing the right visual representation for your data. Depending on the data and the audience, tables, different types of charts, or graphs may be more effective in conveying information. Tailoring your visuals to the content and audience is essential.

3. **Address Readability**: Addressing the issue of plot readability is crucial. Ensuring that your plots are clear, properly labeled, and well-structured is vital for effective communication. Make readability a priority to enhance the audience's understanding of your analysis.

Part 2. Create a model card

	Property	Decision Tree	Naive Bayes	K-Nearest Neighbor	Logistic Regression	Support Vector Machine (SVM)
0	Parametric or Non-Parametric	Non-Parametric	Parametric	Non-Parametric	Parametric	Non-Parametric
1	Input (Continuous, Discrete, Mixed)	Mixed (Both)	Discrete	Continuous (Both)	Continuous (Both)	Continuous (Both)
2	Output (Continuous, Discrete, Mixed)	Discrete	Discrete	Discrete	Discrete	Discrete
3	Can the model handle missing values	Yes	Yes	Yes	Yes	Yes
4	Model Representation	Tree-like structure	Probability tables	Stores instances	Linear Equation	Hyperplanes in multi-dimensional space
5	Model Parameters	Splitting criteria (e.g., Gini, Entropy)	Class probabilities	Number of neighbors, distance metric	Coefficients and intercept	Support vectors, kernel function
6	How to make the model more complex	Increase tree depth or number of nodes	Use more features or feature interactions	Increase k (number of neighbors)	Increase feature dimensions or use polynomial ...	Use a more complex kernel (e.g., RBF)
7	How to make the model less complex	Prune the tree, limit max depth	Use fewer features	Decrease k	Regularization (e.g., L1, L2)	Use a simpler kernel (e.g., linear)
8	Is the model interpretable or transparent	Interpretable, decision rules	Interpretable, conditional probabilities	Semi-interpretable (can visualize neighbors)	Interpretable (coefficients indicate feature i...	Less interpretable due to complex hyperplanes