Subject: 23CSE301

Lab Session: 10

Notes:

1. Please read the assignment notes carefully and comply to the guidelines provided.
2. Code should be checked into the GitHub. These details shall be provided in the Lab.
3. If you have not completed the prerequisite assignments, please complete them before the next lab session.

Coding Instructions:

1. The code should be modularized; The asked functionality should be available as a function. Please create multiple functions if needed. However, all functions should be present within a single code block, if you are using Jupyter or Colab notebooks.
2. There should be no print statement within the function. All print statements should be in the main program.
3. Please use proper naming of variables.
4. For lists, strings and matrices, you may use your input values as appropriate.
5. Please make inline documentation / comments as needed within the code blocks.

Main Section (Mandatory):

Please use the data associated with your own project.

Ref:

https://scikit-learn.org/stable/modules/feature_selection.html

https://scikit-learn.org/1.0/modules/generated/sklearn.feature_selection.SequentialFeatureSelector.html

https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html

https://www.kaggle.com/code/teesoong/explainable-ai-on-a-nlp-lstm-model-with-lime

https://www.kaggle.com/code/bextuychiev/model-explainability-with-shap-only-guide-u-need

https://www.kaggle.com/code/ritzig/classification-feature-selection-shap-tutorial

A1. Please perform feature correlation analysis and use heatmap plot for study of the correlation.

A2. Use PCA for reducing the features. Use cumsum variance to retain 99% of the explained variance. With the PCA transformed features, run the classification models for your project. If your project deals with regression, use regression models. Compare the results of the model performance with reduced dimensionality.

A3. Repeat A2 experiment with 95% of explained variance.

A4. Use sequential feature selection / reduction techniques for feature reduction. Compare the results of various models with results obtained by A2 and A3 experiments.

A5. Use LIME and SHAP for explaining the model behaviour. Study the outputs of LIME and SHAP and compare the explainability and usage suitability for these xAI packages.

Optional Section:

O1. Use other explainable tools for explaining your model behaviour.

Report Assignment:

Please update your last week's report in IEEE format. Expand the methodology and results sections with outcomes of this experiments & results obtained. Please discuss your observations, inferences in results & discussion section. Please conclude the report appropriately with these experiments. Consider following points for observation analysis & inferences. Follow the following instructions:

1. Please tabulate the results (don't take screenshots and add to report)
2. Each table and figure should be captioned and cited in text
3. Refer to the "WritingPaper.pptx" file for more instructions and guidelines