|  |  |
| --- | --- |
| **ASSIGNMENT 2: RESEARCH OUTLINE** | |
| **1) Business Understanding** | |
| 1.1. Thesis (As Question): |  |
| 1.2. Expectation (Yes/No): | No / Yes |
| 1.3. Baseline: | Internal / External, Metrics (ex: R2, MSE, F1-score, Lorenz Curve, Gini): |
| 1.4. Expected Value-Add Metrics: |  |
| **2) Data Understanding** | |
| 2.1. Dataset: | Size (#Rows x #Cells): Source(s): |
| 2.2. List Categorical (CF) & Numerical Features (NF): |  |
| 2.3. Transforming CF/NF into one another: | No / Yes. If yes, techniques: |
| 2.4. Assigning Independent & Dependent Variables: | No / Yes. If yes, techniques:  If yes, list ID & DV: |
| 2.5. Changing Dimensionality: | No / Yes. If yes, techniques: |
| 2.6. Time series: | No / Yes |
| 2.7. List Sampling and Resampling Methods: |  |
| 2.8. List other data manipulation / prep techniques: |  |
| **3) Model Understanding** | |
| 3.1. How are you generating the baseline : | No / Yes. If yes, techniques: |
| 3.2. Are your forecasting: | No / Yes. If yes, time horizon: |
| 3.3. Model Selection (circle or add): | SLR, OLS, Stepwise, Logit, Probit, Anova, CART, Lasso, Boosted, Delphi, Neural, SVM  Others: |
| 3.4. For each technique, list key parameters and settings: |  |
| 3.5. Will you change parameter settings to improve fit: | No / Yes. If yes, list changes: |
| **4) Performance / Fit** | |
| 4.1. Do you expect to outperform on the baseline: | No / Yes |
| 4.2. What fit metrics will you use (circle or add): | R2, pseudo R2, MSE, RMSE, Type 1/2 Error, F1-score, Lorenz Curve, Gini  Others: |