Available From: 05 Jun 25 11:25 AM

Due Date: 30 Jun 25 12:30 AM

Total Marks: 60

Submission Type: File Upload

Description ⌃

Business Context

Renewable energy sources play an increasingly important role in the global energy mix, as the effort to reduce the environmental impact of energy production increases.

Out of all the renewable energy alternatives, wind energy is one of the most developed technologies worldwide. The U.S. Department of Energy has put together a guide to achieving operational efficiency using predictive maintenance practices.

Predictive maintenance uses sensor information and analysis methods to measure and predict degradation and future component capability. The idea behind predictive maintenance is that failure patterns are predictable and if component failure can be predicted accurately and the component is replaced before it fails, the costs of operation and maintenance will be much lower.

The sensors fitted across different machines involved in the process of energy generation collect data related to various environmental factors (temperature, humidity, wind speed, etc.) and additional features related to various parts of the wind turbine (gearbox, tower, blades, break, etc.).

Objective

"ReneWind" is a company working on improving the machinery/processes involved in the production of wind energy using machine learning and has collected data on generator failure of wind turbines using sensors. They have shared a ciphered version of the data, as the data collected through sensors is confidential (the type of data collected varies with companies). Data has 40 predictors, 20000 observations in the training set, and 5000 in the test set.

The objective is to build various classification models, tune them, and find the best one that will help identify failures so that the generators can be repaired before failing/breaking to reduce the overall maintenance cost.
The nature of predictions made by the classification model will translate as follows:

- True positives (TP) are failures correctly predicted by the model. These will result in repair costs.
- False negatives (FN) are real failures where there is no detection by the model. These will result in replacement costs.
- False positives (FP) are detections where there is no failure. These will result in inspection costs.

It is given that the cost of repairing a generator is much less than the cost of replacing it, and the cost of inspection is less than the cost of repair.

"1" in the target variable should be considered as "failure" and "0" represents "No failure".

Data Dictionary

The data provided is a transformed version of the original data which was collected using sensors.

- Train.csv - To be used for training and tuning of models.
- Test.csv - To be used only for testing the performance of the final best model.

Both datasets consist of 40 predictor variables and 1 target variable.

← →

Both datasets consist of 40 predictor variables and 1 target variable.

Submission Guidelines :

1. There are two ways to work on this project:

i. Full-code way: The full code way is to write the solution code from scratch and only submit a final Jupyter notebook with all the insights and observations.

ii. Low-code way. The low-code way is to use an existing solution notebook template to build the solution and then submit a business presentation with insights and recommendations.

The primary purpose of providing these two options is to allow learners to opt for the approach that aligns with their learning aspirations and outcomes. The below table elaborates on these two options.

Submission type	Who should choose	What is the same across the two	What is different between the two	Final submission file [IMP]	Submission Format
Full-code	Learners who aspire to be in hands-on coding roles in the future focussed on building solution codes from scratch	Perform exploratory data analysis to identify insights and recommendations for the problem	Focus on code writing: 10-20% grading on the quality of the final code submitted	Solution notebook from the full-code template submitted in .html format	.html
Low-code	Learners who aspire to be in managerial roles in the future that focus on solution review, interpretation, recommendations, and communicating with business	Focus on business presentation: 10-20% grading on the quality of the final business presentation submitted	Business presentation in .pdf format with problem definition, insights, and recommendations	.pdf	

Please follow the below steps to complete the assessment.Kindly note that if you submit a presentation, ONLY the presentation will be evaluated. Please make sure that all the sections mentioned in the rubric have been covered in your submission.

i. Full-code version

- Download the full-code version of the learner notebook.
- Follow the instructions provided in the notebook to complete the project.
- Clearly write down insights and recommendations for the business problems in the comments.
- Submit only thesolution notebook prepared from the learner notebook [format: .html]

ii. Low-code version

- Download the low-code version of the learner notebook.
- Follow the instructions provided in the notebook to complete the project.
- Prepare a business presentation with insights and recommendations for the business problem.
- Submit only thepresentation [format: .pdf]

2. Any assignment found copied/plagiarized with other submissions will not be graded and awarded zero marks.

3. Please ensure timely submission as any submission post-deadline will not be accepted for evaluation.

4. Submission will not be evaluated-

- If it is submitted post-deadline, or,
- If more than 1 file is submitted.

Best Practices for Full-code submissions

- The final notebook should be well-documented, with inline comments explaining the functionality of code and markdown cells containing comments on the observations and insights.
- The notebook should be run from start to finish sequentially before submission.
- It is important to remove all warnings and errors before submission.
- The notebook should be submitted as an HTML file (.html) and NOT as a notebook file (.ipynb).
- Please refer to the FAQ page for common project-related queries.

Best Practices for Low-code Submissions

- The presentation should be made with the audience being the company's Data Science lead in mind.
- The key points in the presentation should be the following:

  - Business Overview of the problem and solution approach
  - Key findings and insights that can drive business decisions
  - Business recommendations
  - Focus on explaining the key takeaways in an easy-to-understand manner.
  - Including the potential benefits of implementing the solution will give you the edge.

- Copying and pasting from the notebook is not a good idea, and codes should not be shown unless they are the focal point of your presentation.
- The presentation should be submitted as a PDF file (.pdf) and NOT as a .pptx file.
- Please refer to the FAQ page for common project-related queries.

Power Ahead!

Rubric

Criteria

Exploratory Data Analysis

- Data Overview - Univariate Analysis - Bivariate Analysis

Points

5

Criteria

Data Preprocessing

Criteria

Model Building

- Choose the metric of choice with a rationale - Build a Neural Network model with SGD as the optimizer - Comment on model performance

Points

6

Criteria

Model Performance Improvement and Final Model Selection

- Use combinations (at least 6) of different methods mentioned below to improve the model performance - More hidden layers - Different optimizers (SGD, Adam) - Dropout - Class Weights - Comment on model performance for each of the models - Choose the best model from the ones built with proper reasoning

Points

34

Criteria

Actionable Insights & Recommendations

Conclude with the key takeaways (actionable insights and recommendations) for the business

Points

3

Criteria

Presentation/Notebook - Overall Quality

- Structure and flow - Crispness - Visual appeal - Conclusion and Business Recommendations OR - Structure and flow - Well commented code - Conclusion and Business Recommendations

Points

8

Submit Assignment ⌃

📄 Submission Type

File Upload

Upload File ⬆
Type: HTML PDF • Maximum File Size: 100 MB

🎧 Comments