**Capstone Report Summary: Predicting Technician Actions from Work Order Descriptions**

**Objective:**The goal of this project was to use past work order data to see if it was possible to teach a model to understand the relationship between a short description of an issue and the technician's written notes or solution. Ideally, this would act as a framework to aid in the creation of a maintenance application that can suggest solutions for future trouble calls based on user input of the machine breakdown

**Dataset:**

* A spreadsheet of 70,000+ work orders, cleaned to 40,000+
* Each row has a short Description of the problem. A Text field where the technician wrote what they did to fix or solve the issue. A Work Order Number to track, and machine or equipment number (didn’t use them this time).

**Cleaning the Data**

* Removed empty and junk rows
* Cleaned up spaces, punctuation, and 'nan' strings
* Created two new columns: Description\_cleaned, Text\_cleaned
* The raw data CSV file needed cleaned manually yet, data embedded technician names in the notes, which is why I think my results were as close to what I wanted as it should have been.

**Grouping Notes by Behavior**

* Used TF-IDF to turn the Text\_cleaned into numbers the model could understand
* Ran KMeans clustering (k=5) to group technician responses into 5 clusters
* This helped to assign labels like "reset", "inspect", etc.

**Training a Model**

* Used Description\_cleaned as input (X) and Note\_Cluster as target (y)
* Created a pipeline with TF-IDF vectorization and logistic regression
* Data split 80/20 for training and testing
* Achieved best performance with up to 0.74 accuracy
* Tried different data-cleaning iterations that produced results between 0.67 and 0.74
* Final classification report included precision, recall, and F1-score for each cluster

**Confusion Matrix + Heatmap**

* Generated and plotted to visualize prediction distribution
* Showed the model performed strongest on most common clusters

**TF-IDF Cosine Similarity for Search Function**

* Additional feature added to allow entering a description and returning similar technician notes
* Used cosine similarity between TF-IDF vectors of descriptions

**Performance Evaluation:**

Accuracy: 0.74

Weighted F1 Score: 0.67

Cluster 0: Precision = 0.76, Recall = 0.96 (Most Common)

Cluster 1: Precision = 0.47, Recall = 0.19

Cluster 2: Precision = 1.00, Recall = 0.01

Cluster 3: Precision = 0.38, Recall = 0.05

Cluster 4: Precision = 0.27, Recall = 0.03

* The model did great on the most common tasks (like reset), but wasn’t as accurate for the rare ones such as repaired, fixed, changed, replaced.
* To better visualize how predictions match actual outcomes, a confusion matrix heatmap was added. This helped to see that while the model performs very well on the most frequent cluster, there is some confusion between adjacent categories.

**Prediction Function**

* A custom function was created to test inputs
* Returns predicted cluster and closest matching past technician notes from the same group

**Issues/Considerations and Future Improvements**

* Technician names embedded in notes reduce quality
* Some labels like "completed" are non-informative and reduce training quality
* Cluster imbalance affects recall on minority groups
* Consider fine-tuning the model using embeddings (e.g., MiniLM)
* Improve note preprocessing with domain-specific filters (e.g., keep 'Allen Bradley' but remove 'brad')
* Use true label classification in future if data is annotated manually

John Fizer