Computational Gastronomy: Recipe Review Sentiment Analysis

Independent Project,
Complex Systems Laboratory, IIITD

Mentor: Dr. Ganesh Bagler

Weekly Reports

Week -1

Summary of meeting on 7th Jan 2025

1. Discussion on the dataset, which contains 11 Lakh entries, discussion on how to deal with imbalance using SMOTE, TOMEK links, and challenges.
2. Discussed how we can find unique recipes using recipe IDs to train the model, taking the mean of ratings, and using vector embeddings for review.
3. We discussed the results of the Logistic Regression (78% acc) and Random Forest (60% acc) models.

TO DO:

1. Apply techniques to deal with data imbalance.

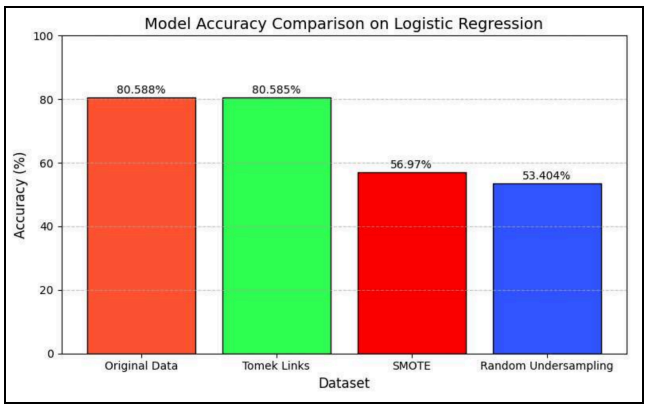
Week -2

Summary of meeting on 14th Jan 2025

1. Presenting the results after SMOTE, tomek links, and random undersampling.

Will attach the summarising graph showing accuracies. **TO DO:**

1. Apply Lazy classifier.
2. Build more models like XGboost, RF etc



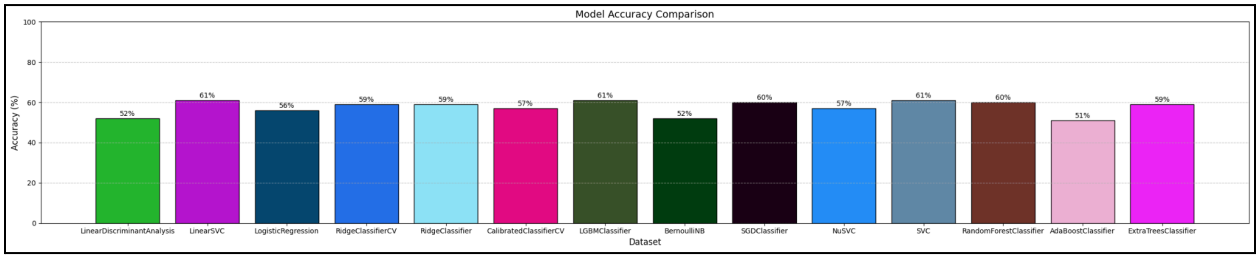
Week - 3

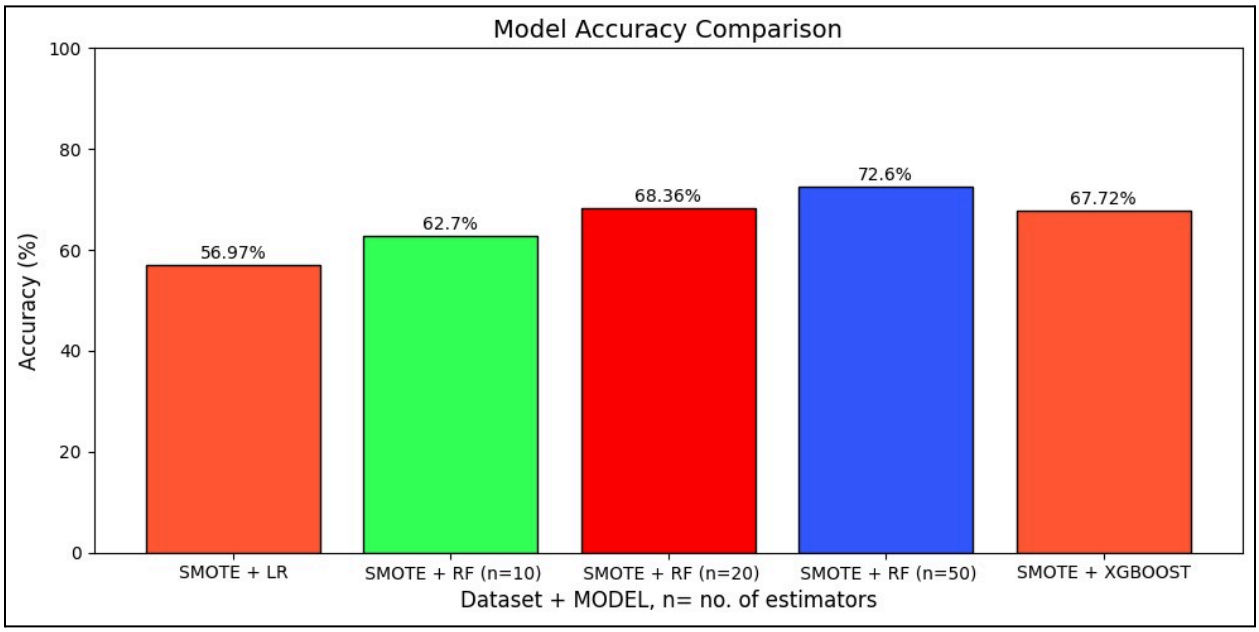
Summary of meeting on 21st Jan 2025

1. Discussed the results of RF and XGBoost on the SMOTE data.
2. Discussed the results of LazyClassifier on a small subset (10000 samples per rating)
3. Discussed with Vibhuti Sir regarding challenges faced (compute power, data imbalance)

Suggestions: Use doc2vec instead of word2vec, not take average rating. To Do:

1. Use Doc2Vec, optimise classical ML models, and run on various advanced models.



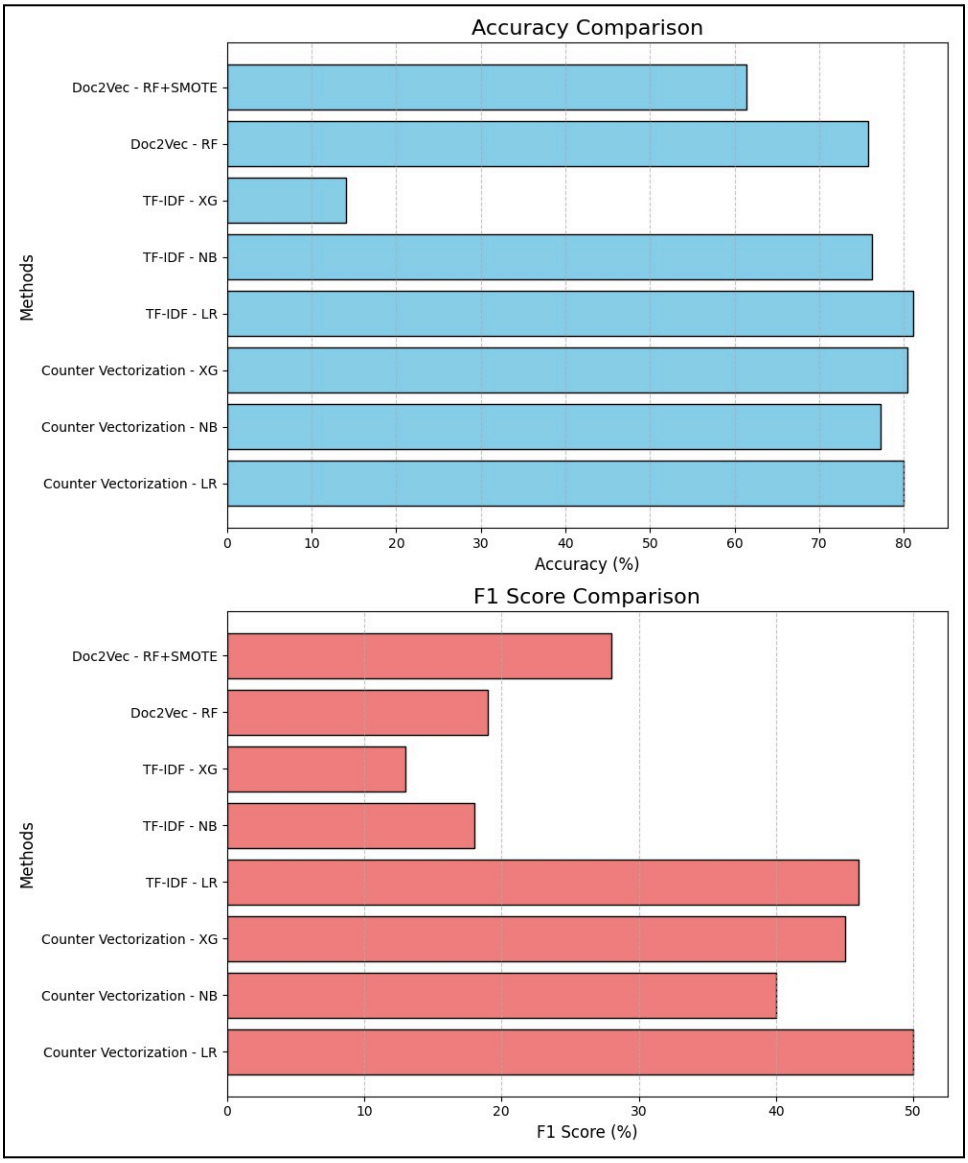


Week 4

Summary of work done , 28th Jan 2025

1. As per Vibhuti Sir's suggestion we used Doc2Vec.
2. In Addition to Doc2Vec, we used different word embedding strategies we employed : a. TF-IDF b. Counter Vectorisation

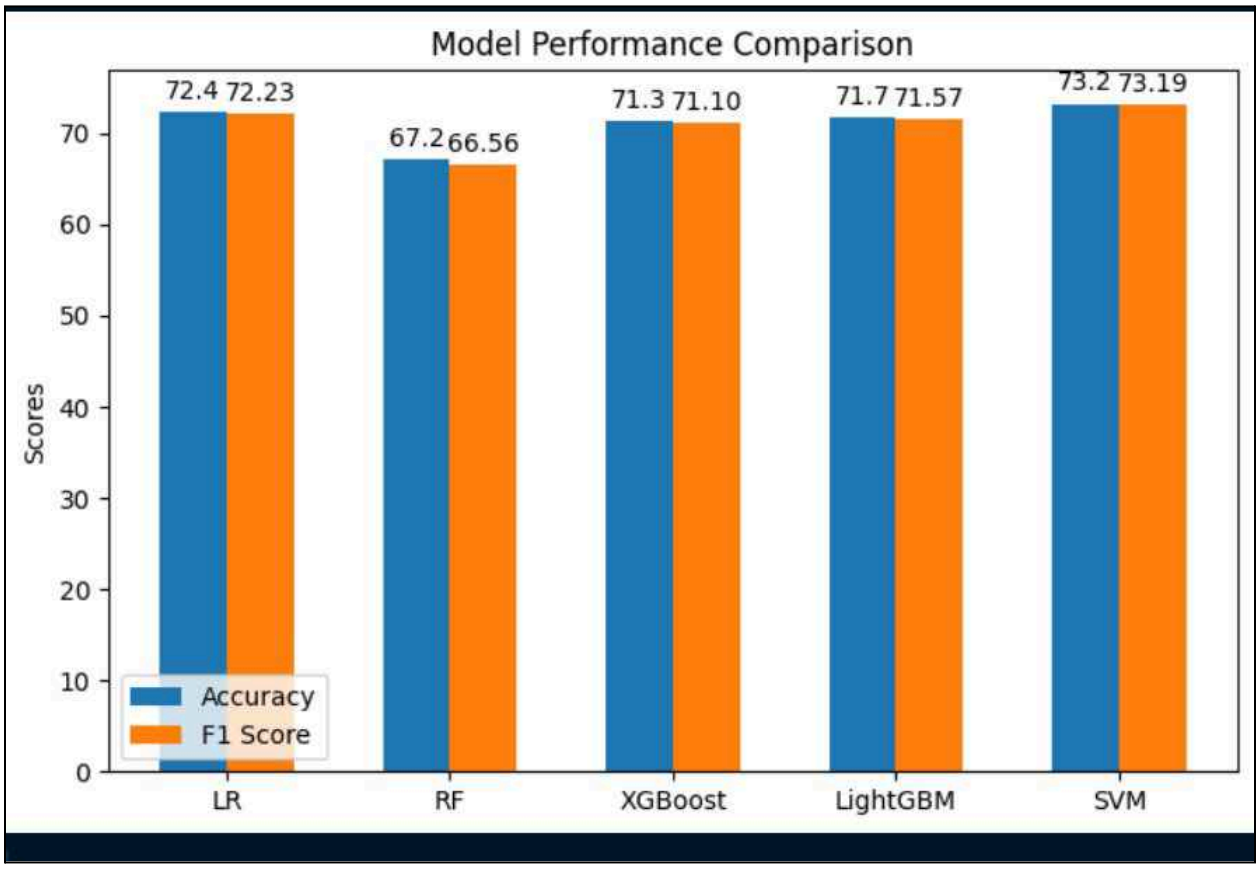
Accuracies on different models (RF, LR, NB, XGboost) and F1 score graph are attached below. We conclude that Doc2vec showed no improvement in computation load and accuracy.



Week 5

Summary of Work Done till 4th Feb 2025

Removed data imbalance by grouping the ratings, i.e. grouped 1 and 2 rating to poor, 3 to neutral, and 4,5 to positive, then undersampling the majority classes to the count of the minority class. The above steps solved the problem of data imbalance, with no need to use synthetic samples, and F1 scores increased to 72% from 50% earlier (without smote) and (22 % with SMOTE). We evaluated different models on this dataset; we trained LR, RF, XGBoost, LGBM, SVMC, and RNNs. We tried to do hyperparameter tuning and also got the best results, 73.18% and the highest F1 score of 73. We achieved the highest accuracy on RNNs. Please find attached the graphs.

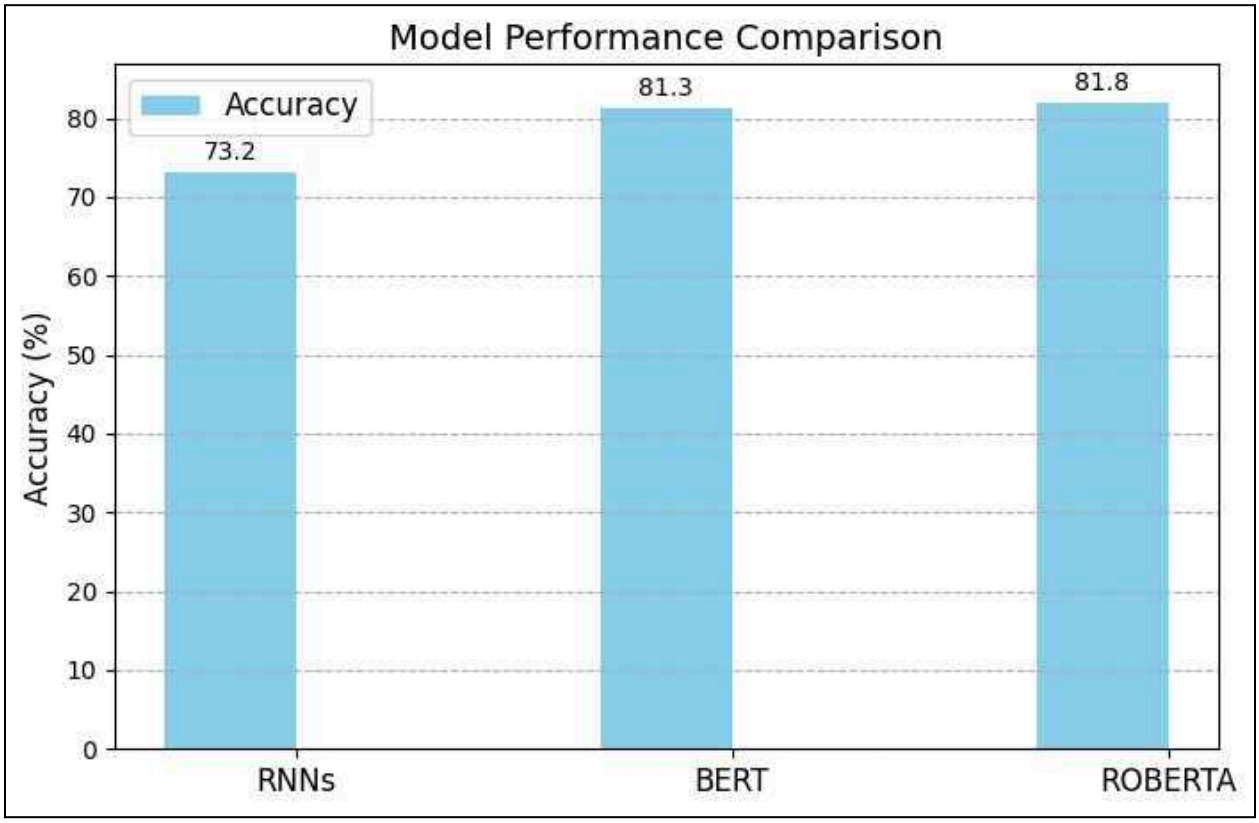


Week 6

Summary of Work Done as of February 11, 2025 We employed advanced NLP models, specifically BERT and RoBERTa, for sentiment analysis, achieving accuracies of 81.3% and 81.8%, respectively. In comparison, our earlier models utilizing RNNs attained an accuracy of 73.18%. This represents an improvement of approximately 8% in model performance.

Meeting Minutes Time : 3:10 PM Date : 11th February, 2025

- Discussed the results of BERT and RoBERTa implementation.
- To consider additional evaluation metrics such as specificity, sensitivity, precision, recall, and F1 score for better analysis.
- Save the trained models for future comparisons and reuse.
- To create a table consolidating results from all implemented models.
- Train and evaluate additional models for further improvements.



Week 7

Meeting Minutes Weekly Discussion

Date : 18th February, 2025

1. Discussed the results (accuracies and other metrics) of models we have implemented on the undersampled data.
2. To implement the best performing model on the whole data.