RESULTS OBTAINED

CatBoost

- Program Accuracy (31%): CatBoost predicts program names moderately well, with 31% accuracy. While it isn't highly accurate, it manages to capture some program-specific patterns.
- Degree Accuracy (84%): For degree predictions, CatBoost performs strongly, achieving 84% accuracy. This indicates it is very reliable when it comes to identifying degree-related data.
- Institute Accuracy (81%): The model predicts institute names with 81% accuracy, showing that it is effective at recognizing institute-specific trends.
- Mean Accuracy (58%): Overall, CatBoost has a mean accuracy of 58%, which reflects decent performance. However, its lower accuracy for programs slightly affects its overall effectiveness.

ANN (Artificial Neural Network)

- Program Accuracy (20%): The ANN model struggles significantly to predict program names, with only 20% accuracy. This shows it has difficulty identifying patterns in program data.
- Degree Accuracy (81%): Similar to CatBoost, ANN does well in predicting degree-related data, with an accuracy of 81%.
- Institute Accuracy (19%): The model performs poorly when it comes to predicting institutes, with just 19% accuracy. This is one of its major weaknesses.
- Mean Accuracy (40%): Overall, ANN has a mean accuracy of 40%, making it the weakest model. Its poor performance in predicting programs and institutes contributes to its low overall accuracy.

Random Forest

- Program Accuracy (76%): Random Forest is excellent at predicting program names, with an accuracy of 76%. This shows it effectively captures patterns related to programs.
- Degree Accuracy (94%): The model achieves a remarkable 94% accuracy in degree predictions, making it the most reliable model for this category.
- Institute Accuracy (75%): Random Forest performs well in predicting institutes, with an accuracy of 75%. While slightly lower than CatBoost, it still shows strong reliability.
- Mean Accuracy (81%): With an overall accuracy of 81%, Random Forest is the best-performing model. It is consistent and reliable across all aspects of the dataset.

LSTM Application

The predicted opening and closing ranks in Table 1 were generated using an LSTM model, applied to a dataset filtered under specific conditions. These conditions include for example the institute being IIT-Bombay, the program being Electrical Engineering, the quota set to AI (All India), the pool being Gender-Neutral, the program duration of 4 years, and the category being General (GEN), as shown in Fig. 5. These predictions provide valuable insights into rank trends for this program and candidate pool from 2022 to 2025.

Predicted Opening and Closing Ranks (2022–2025)

Year	Opening Rank	Closing Rank
2022	99.88	435.74
2023	101.39	440.23
2024	103.54	446.45
2025	106.51	454.82

This prediction approach is highly adaptable, allowing users to customize the filtering criteria based on their specific requirements. Candidates can adjust variables such as institute, program, quota, gender pool, program duration, or category to generate tailored predictions for their individual preferences. This flexibility ensures that the model can provide personalized insights for diverse scenarios.

The original data for opening and closing ranks for the IIT-Bombay Electrical Engineering program (AI Quota, Gender-Neutral, 4 Years, General Category) was sourced from **College Pravesh** for the years 2022, 2023, and 2024. The ranks are as follows:
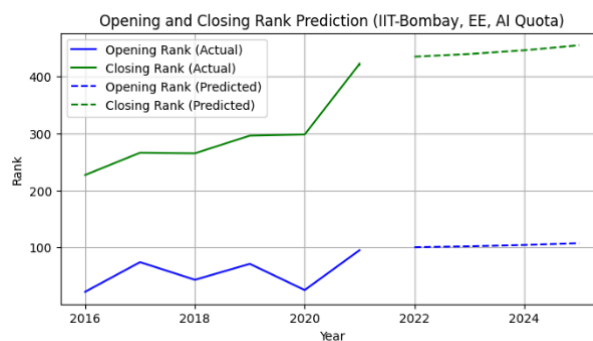
Actual Opening and Closing Ranks (2022–2024)

Year	Original Opening Rank	Original Closing Rank
2022	103	430
2023	142	463
2024	15	464

The LSTM model's predictions closely align with the original recorded data, demonstrating its accuracy and reliability in capturing rank trends.

Additionally, a comparison of the models—CatBoost, ANN, and Random Forest—is shown in Table II. These models were evaluated based on their accuracy in predicting key components like program, degree, and institute:

Model Name	Programme	Degree	Institute	Mean Accuracy
CatBoost	0.31	0.84	0.81	0.58
ANN	0.20	0.81	0.19	0.40
Random Forest	0.76	0.94	0.75	0.81

TABLE II: Comparison of Models

The **Random Forest** model achieved the highest mean accuracy of 81%, proving to be the most reliable model for predicting program, degree, and institute data. Meanwhile, the **LSTM model** was particularly effective in forecasting rank trends, as its predictions were closely aligned with the original data. Together, these models offer accurate and insightful predictions for college programs and rank trends.