

Date: April 12th, 2024

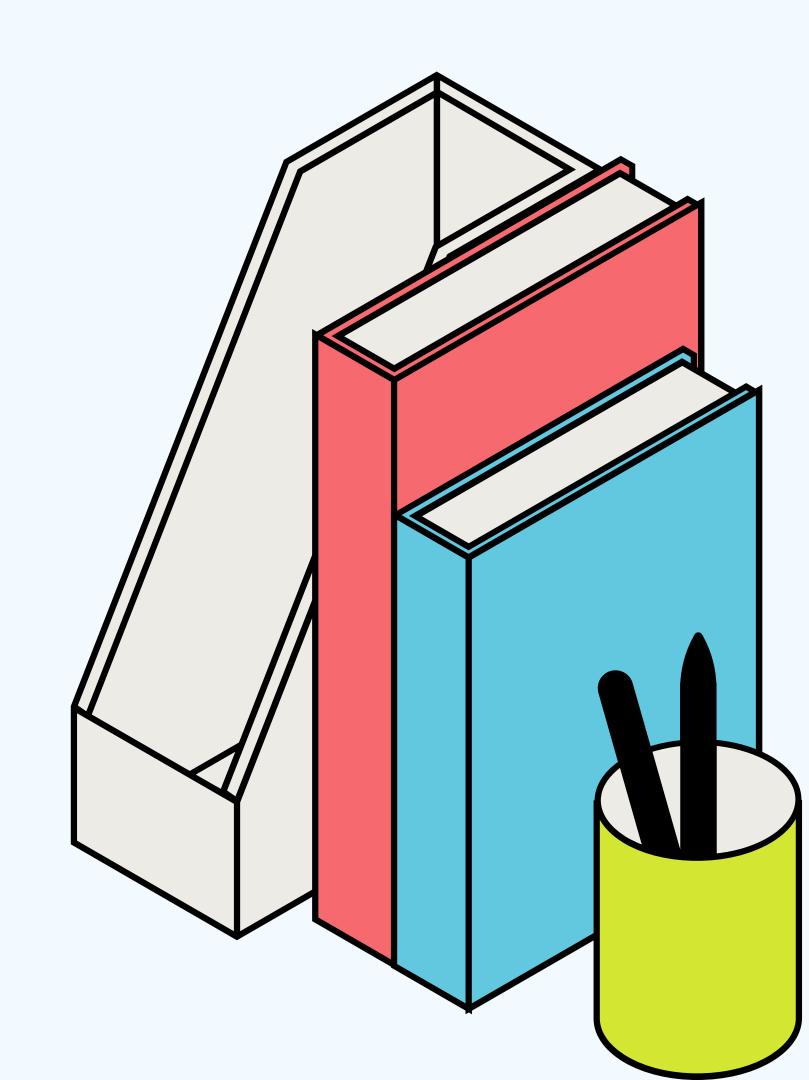
Presented by: JJ Park

BrainStation Data Science Bootcamp



Table of Contents

- 1. Introduction
- 2. Problem Space
- 3. Solution
- 4. Impact
- 5. EDA
- 6. Preprocessing
- 7. Modeling
- 8. Summary





"900M+ Members Worldwide"

"20 Million Job Postings Every Month"

"140 job applications submitted every second"



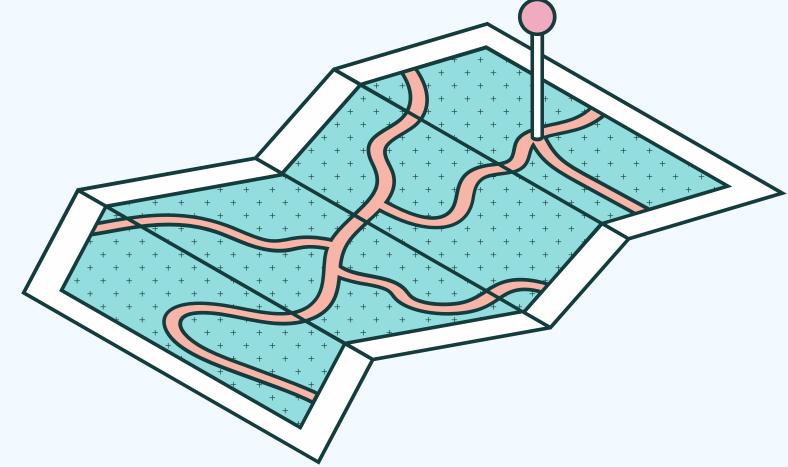
Problem Space

Analyze the 'views' metric to uncover relationships among various job attributes within the dataset.

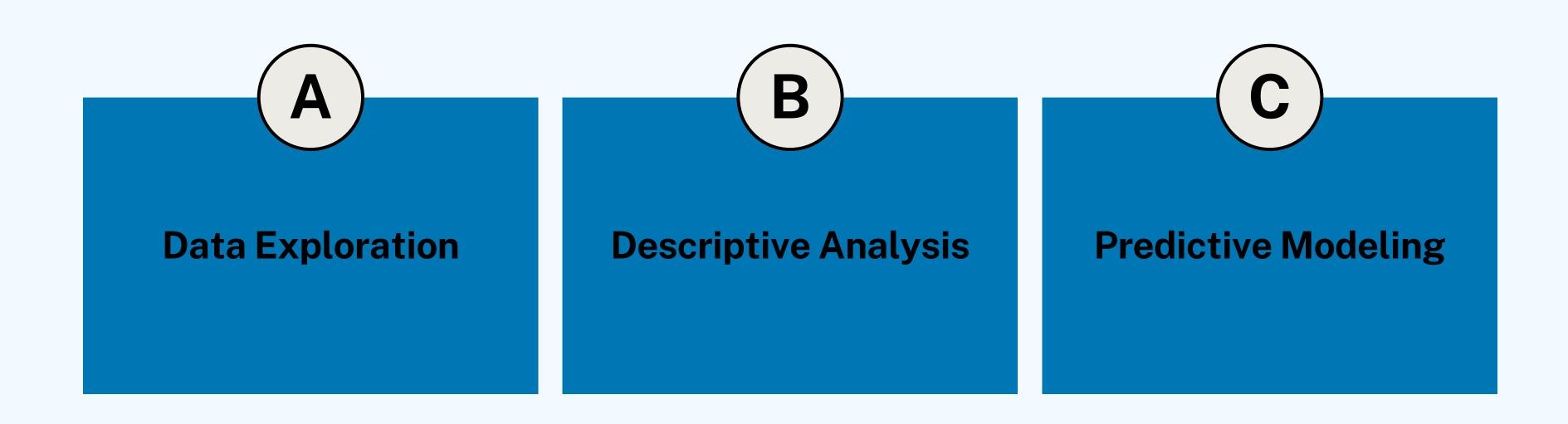
Target Variable: Number of Views

Key Variables: Title, Listed Time, Application, Work

Type, Experience Level



Data Science Solution



Impact

- Hiring Impact: Every minute, 6 people are hired through LinkedIn
- Economic Influence: LinkedIn ads reach approximately 6% of the global population
- Business Leads: About 80% of B2B leads generated through the platform

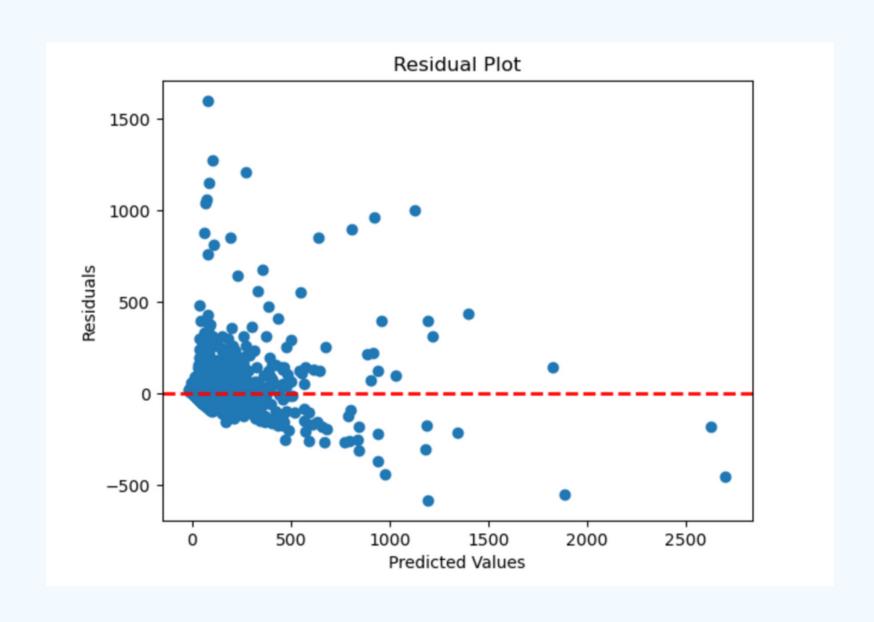
AFFECTS BOTH RECRUITERS AND JOB SEEKERS

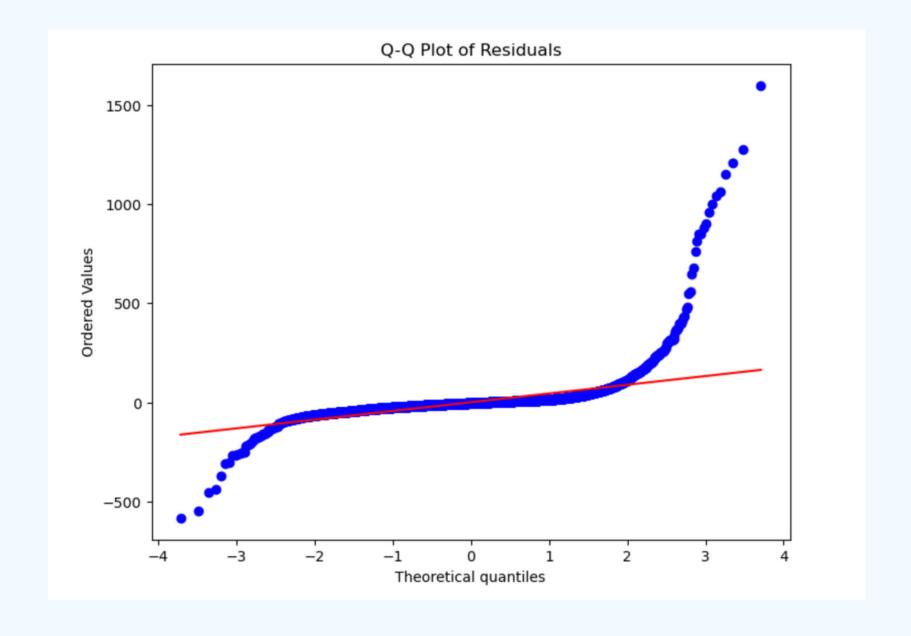
Companies can optimize their job postings and recruitment strategies

Job applicants can improve their chances of securing relevant job opportunities



Linear Regression

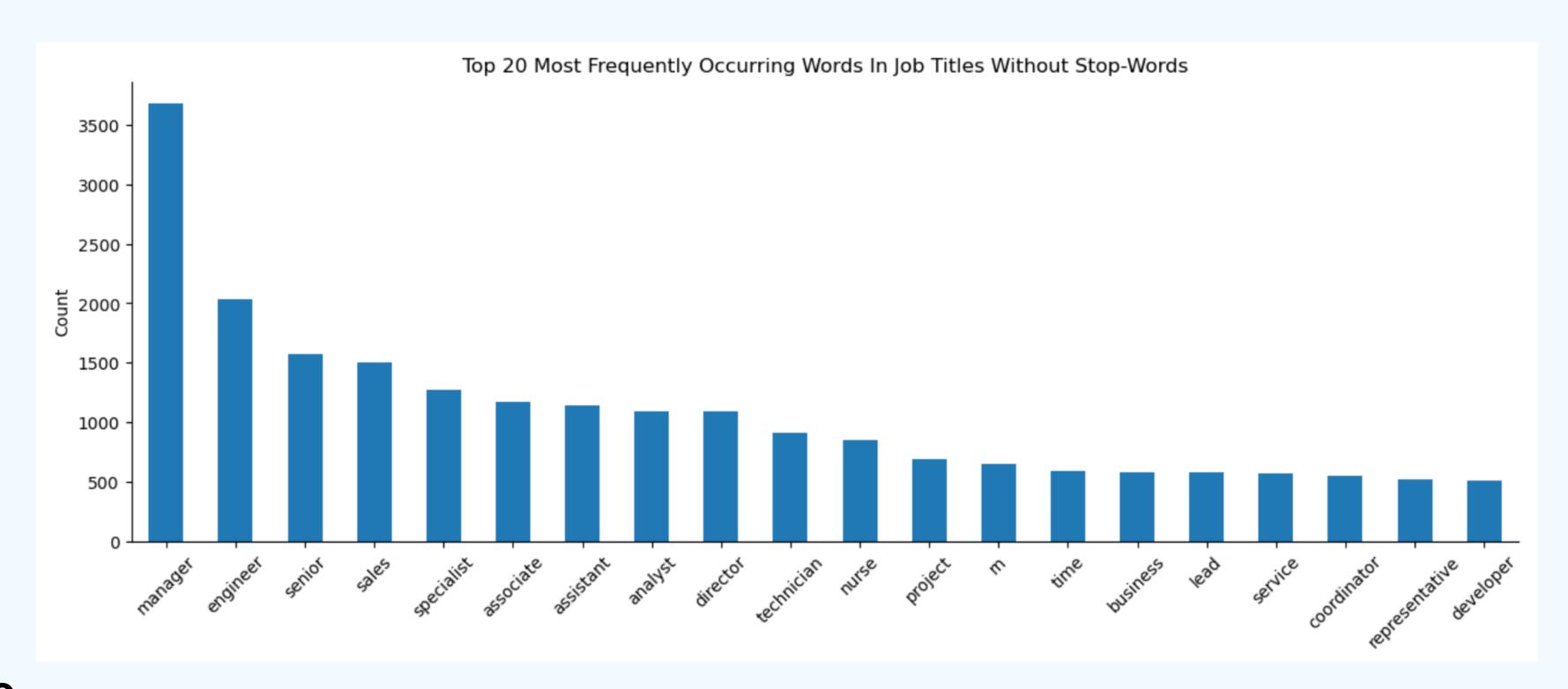




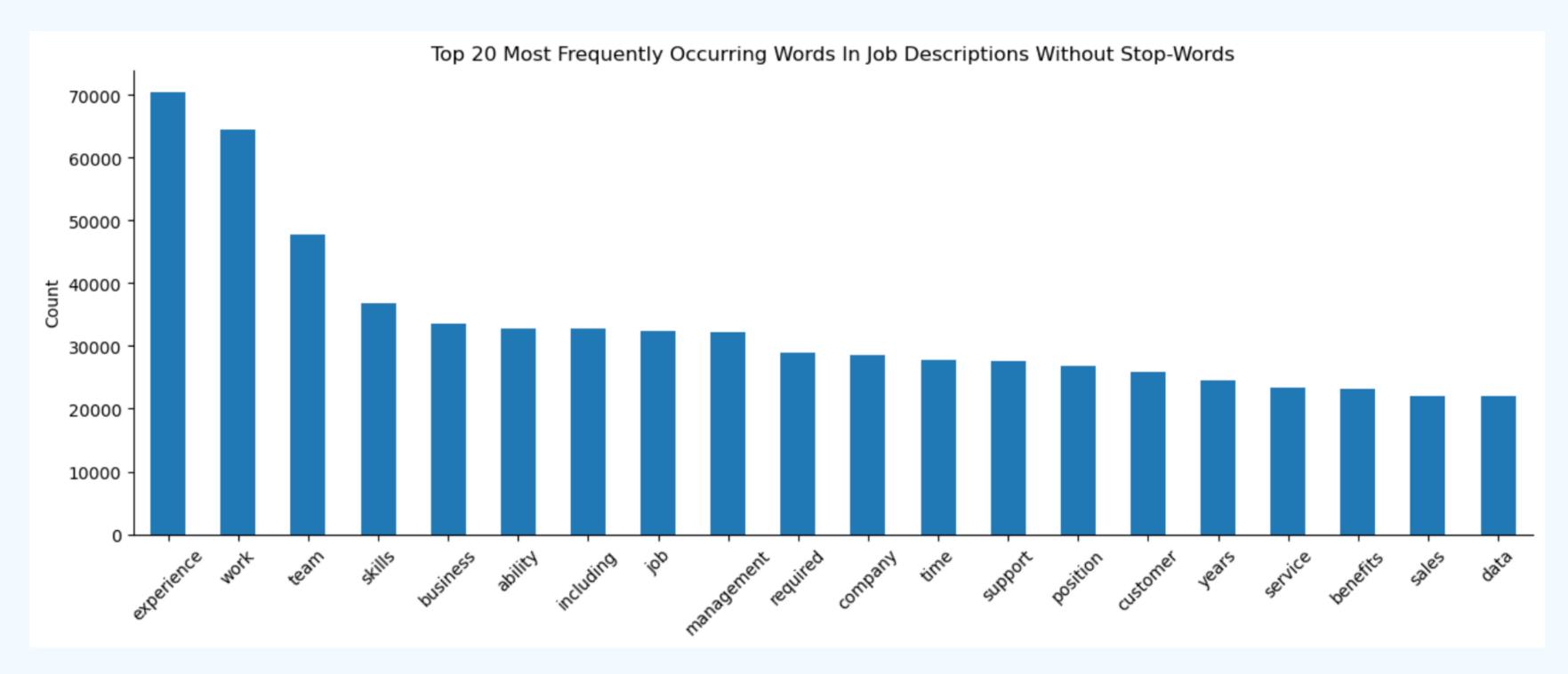
Baseline Model Evaluation Metrics

25.70	25.75	0.7391	0.7126
7.49	19.72	0.9633	0.7674
13.19	20.65	0.9594	0.7116
	7.49	7.49 19.72	7.49 19.72 0.9633

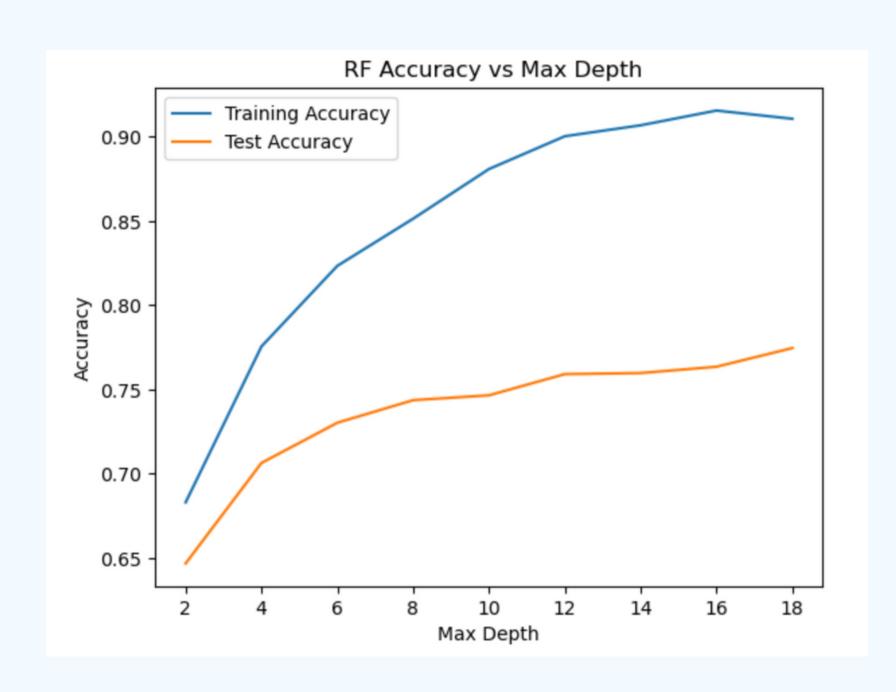
NLP - Title

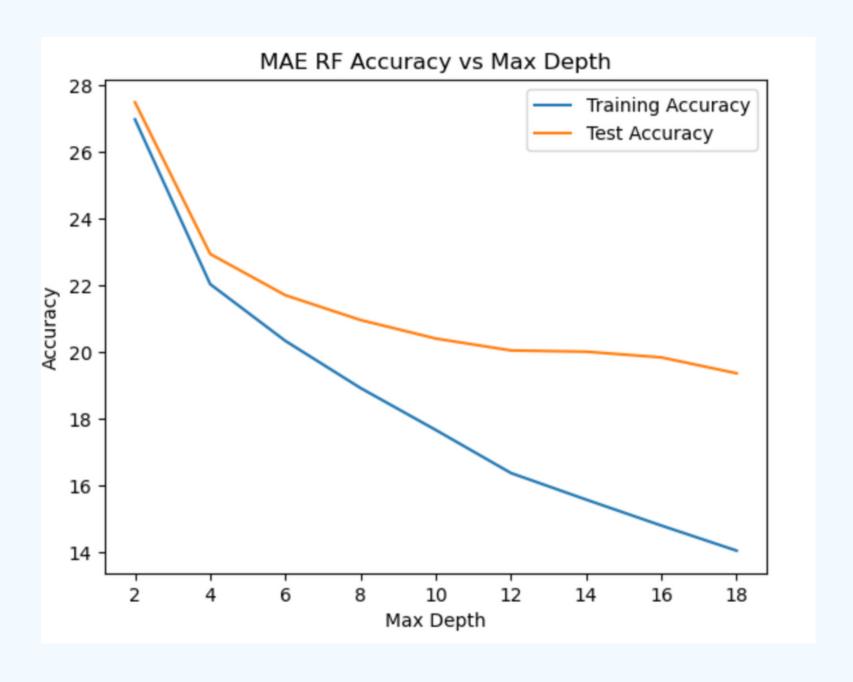


NLP - Description



Hyperparameter Tuning





Random Forest Regressor

Model	MAE (Train)	MAE (Test)	R _^ 2 (Train)	R ₂ (Test)
Random Forest	16.65	20.26	0.9177	0.7570

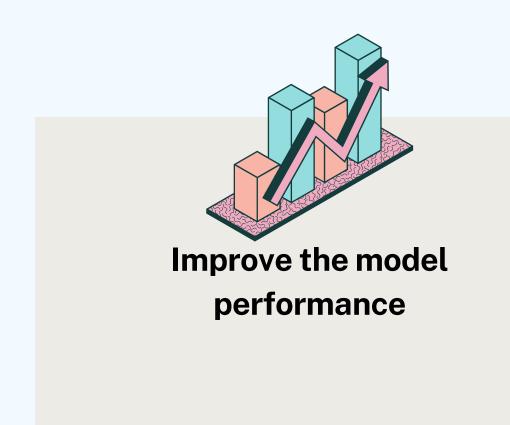
XGBoost Regressor

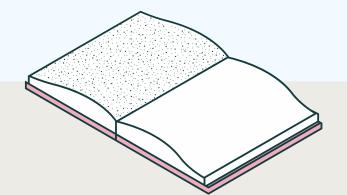
Model	MAE (Train)	MAE (Test)	R ₂ (Train)	R ₂ (Test)
XGBoost	12.44	19.86	0.9695	0.7337

Evaluation Metrics

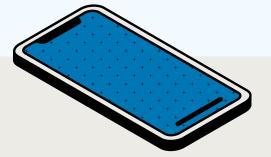
Model	MAE (Train)	MAE (Test)	R-squared (Train)	R-squared (Test)
Random Forest Regressor	16.65	20.26	0.9177	0.7570
XGBoost Regressor	12.44	19.86	0.9695	0.7337

Next Steps





Advanced Text Processing



Generate interactive UI

Thank you for Listening.

