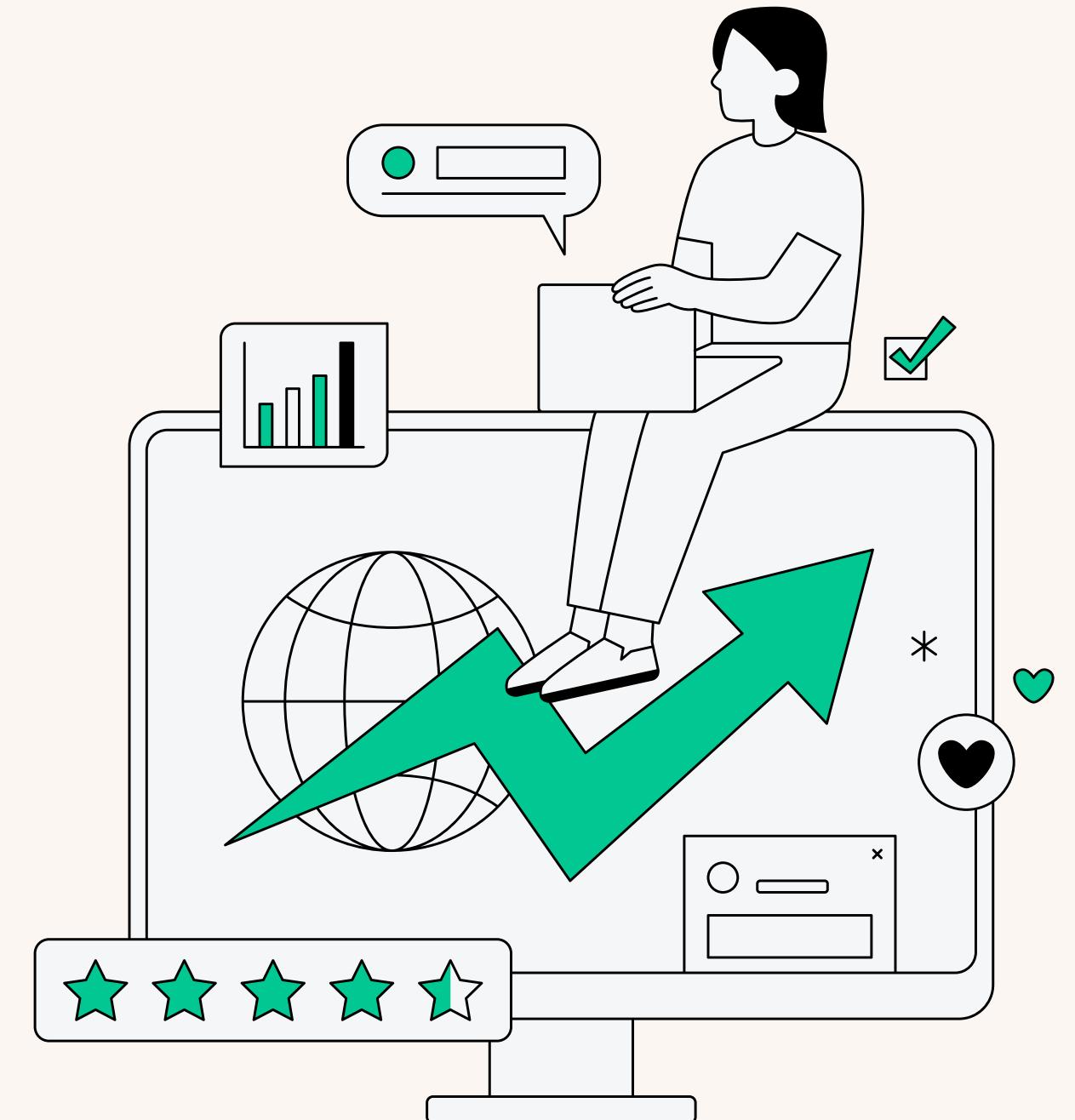


Home valuation Models for Faster Bank Loan Approval

A Machine Learning Solution for Fast and Accurate Home Valuation

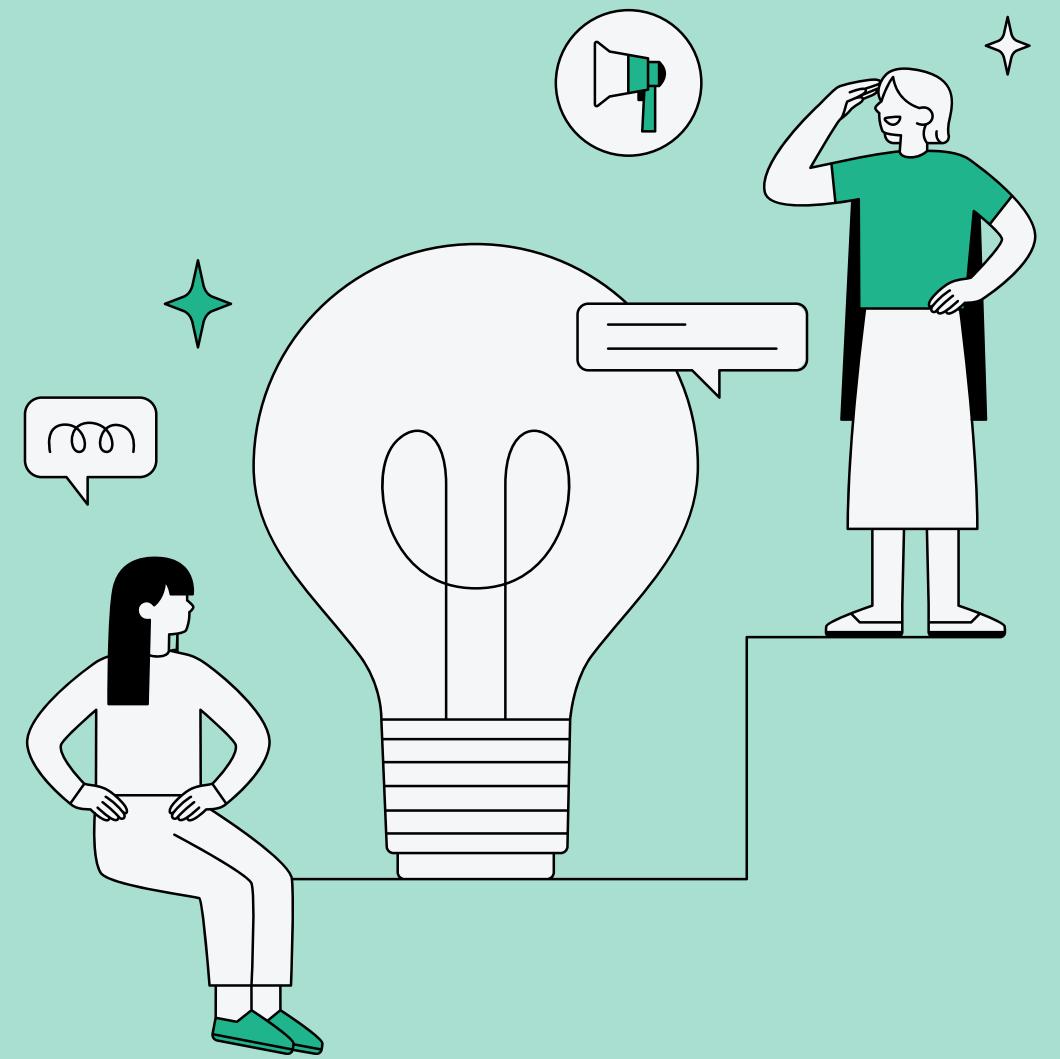
Presented by Elie Niringiyimana

1



Business Problem

- Banks and financial institutions need a reliable and swift way to assess home values for loan approval decisions.
- A fast, data-driven model is required to provide accurate home valuations, ensuring faster loan processing and reducing the risk of overvaluation or undervaluation.



Solution

- Implement a Machine Learning model that classify and predicts home prices based on key property features: location(longitude and latitude), size (lot acres, square feet), number of bedrooms, number of bathrooms, garage, and fireplaces.
- The model will use historical data to learn patterns and provide instant valuations that are data-backed.



Data Understanding

- The original dataset consist of 5000 rows and 16 columns.
- Cleaned dataset contains 14880 house records with 2202 features. The increase in rows was caused by oversampling to fight class imbalance while increase in columns was caused by encoding.

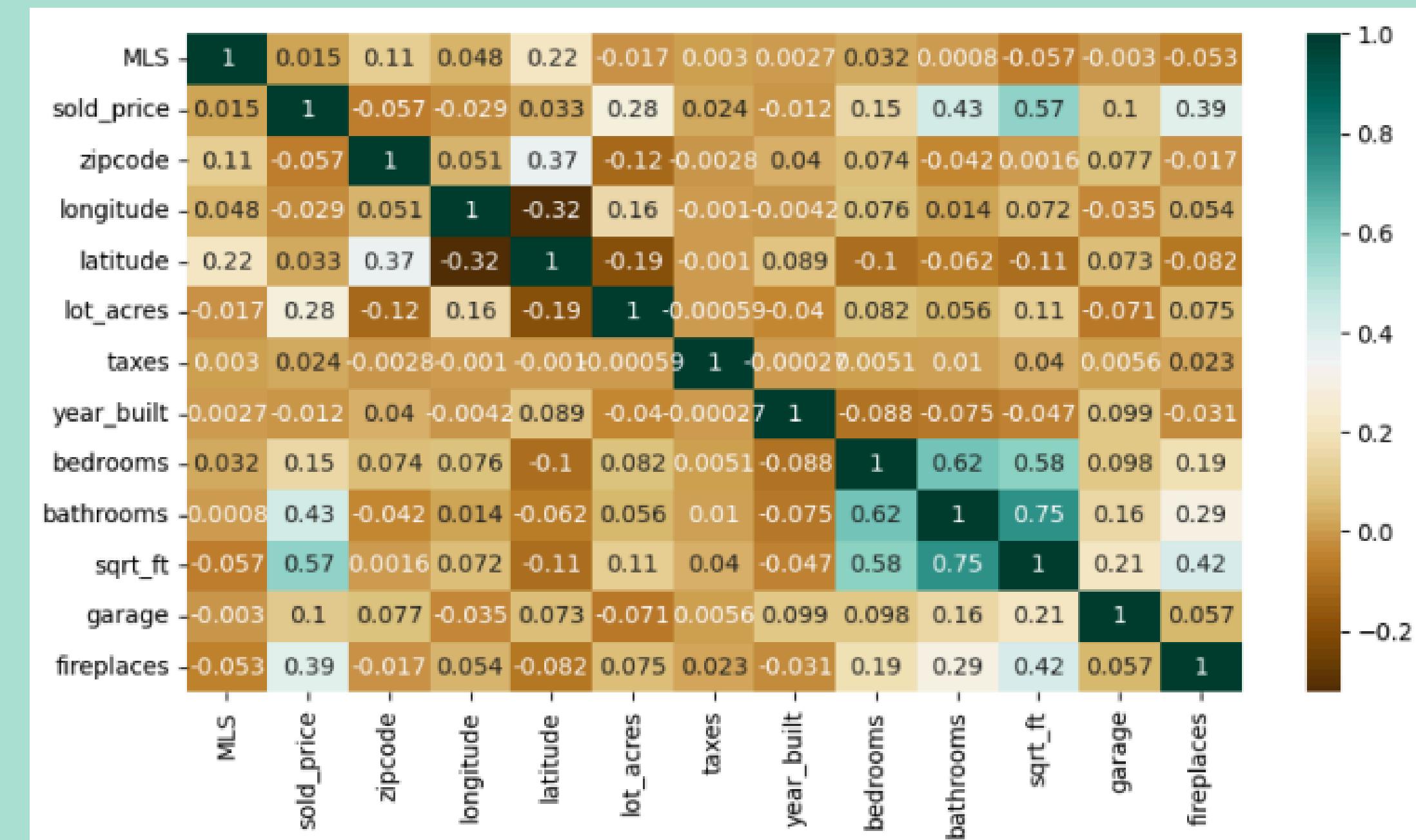
Data Preparation

- Handling Missing Data: Imputed missing values for numerical and categorical features.
- Variable transformation to make data approximately normal
- Encoding Categorical Variables: Converted non-numerical data into usable numerical format.
- Normalization: Standardized numerical features to ensure uniformity and improve model performance.
- Outlier detection and removal
- Feature Selection: Identified key features that significantly influence price predictions.



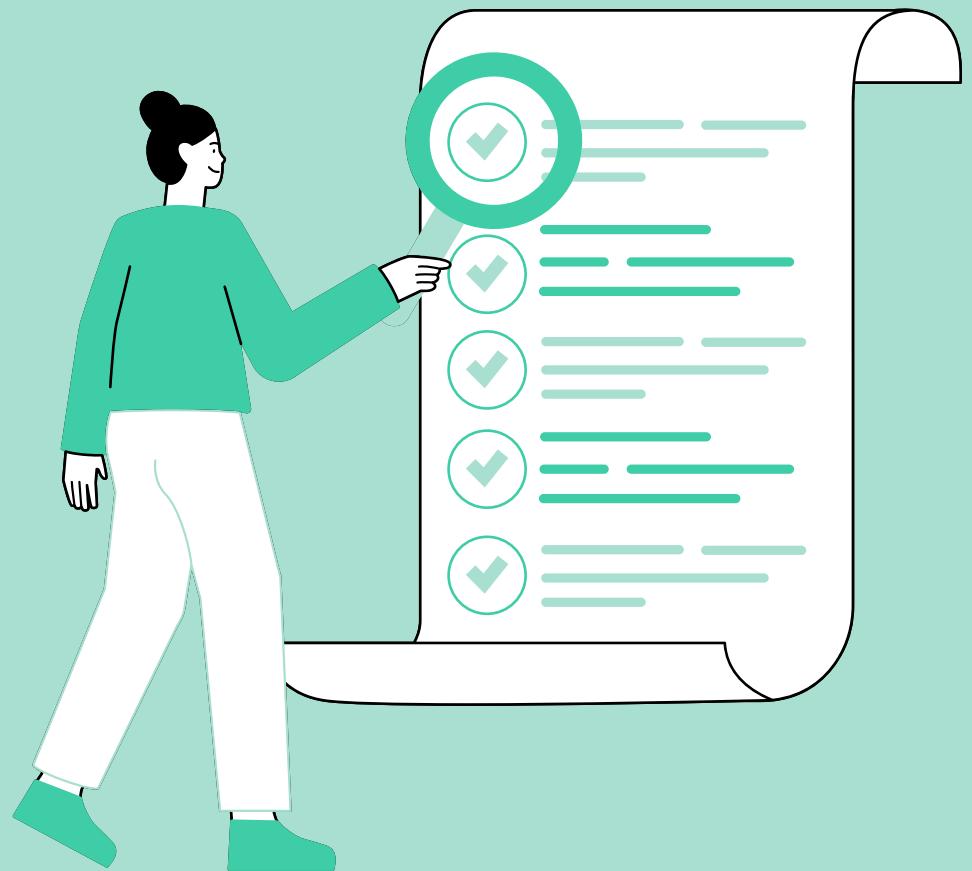
Correlation

Key variables such as lot size (acres), square footage, number of bedrooms, garage, fire places, and number of bathrooms show a correlation with the sale price, suggesting that, when combined with longitude and latitude, we can effectively create model to classify and predict the sales price.



Tasks

- Classify house in one of 3 categories Economy, Standard, and Luxury.
- Predict house price to ease house valuation



Modelling

- For classification task Naive Bayes, Gaussian Naive Bayes, KNN Classifier
- For regression tasks, I used Multivariate linear regression, and KNN regressor

Classification tasks	Parameters	Accuracy	Regression tasks	Parameter	MAPE
Naive Bayes	epsilon = 1e-3	0.703	Multivariate linear regression	eta=0.0005, epochs=4000	45.37
Gaussian Naive Bayes	epsilon=1e-3	0.717	KNN regressor	k=4, epsilon = 1e-3	5.346
KNN Classifier	k=3,epsilon=1e-3	0.998			

Business Value & Benefits

- Faster Loan Approval Process: Speeds up decision-making by providing instant property valuations.
- Improved Risk Management: More accurate price predictions reduce the risk of approving loans on overvalued properties.
- Cost Efficiency: Banks save costs by minimizing the need for third-party appraisals and manual assessments.
- Customer Satisfaction: Faster loan approvals lead to better customer experience and increased customer retention.



Summary

- Problem: Banks need fast, accurate home valuations for loan approvals.
- Solution: Machine learning models classify, and predict house prices based on property features.
- Key Benefits: Faster loan approvals, Reduced manual appraisals, Minimized risk of valuation errors
- Recommendation: Use KNN classifier to classify house category and KNN regressor to calculate house value.
- Conclusion: This model enhances efficiency, accuracy, and customer satisfaction in the loan approval process.

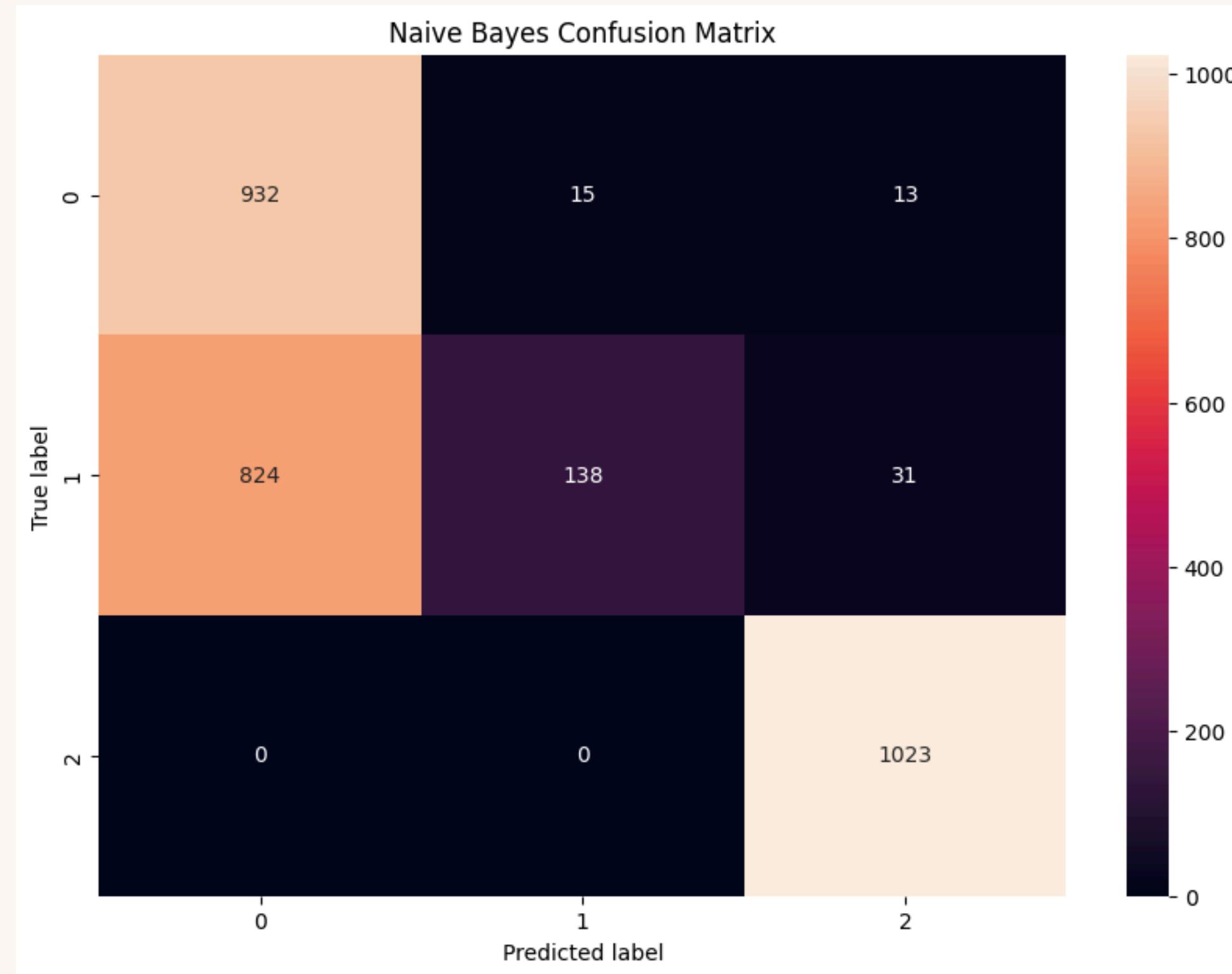
Thank you very much!

Presented by Elie Niringiyimana

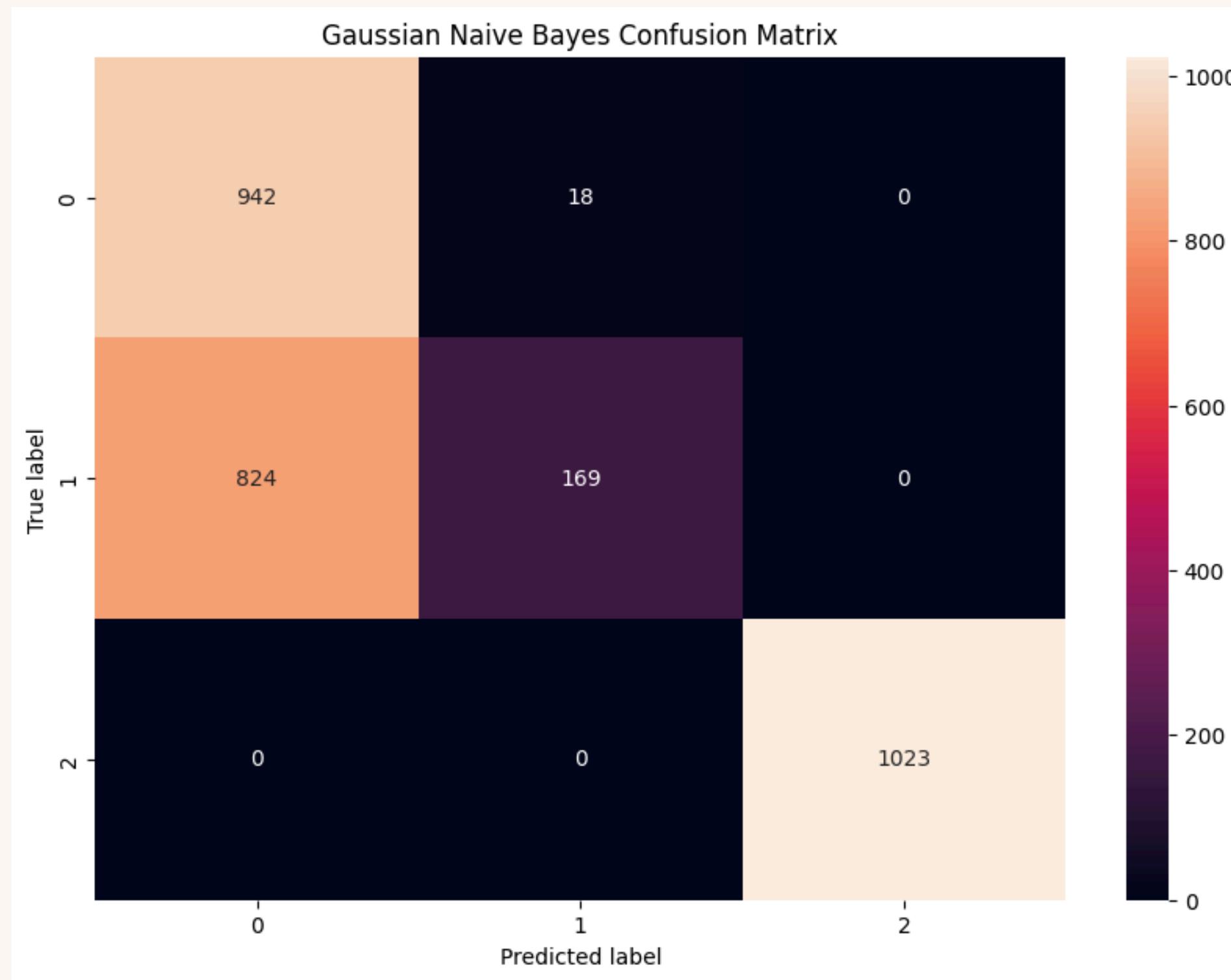
11



Appendix



Appendix



Appendix

