

Detailed Project Report: Author Prediction System for Avila Bible

Contents

1. Introduction
 2. Objectives
 3. Scope
 4. Methodology
 - 4.1 Data Collection and Preprocessing
 - 4.2 Machine Learning Modeling
 - 4.3 User Interaction
 - 4.4 Deployment
 5. Technology Stack
 6. Results and Evaluation
 7. Conclusion
 8. Future Work
 9. References
-

1. Introduction

This Detailed Project Report (DPR) outlines the development of an Author Prediction System for the Avila Bible. The system utilizes machine learning to predict the authorship of text passages using predefined features. The project addresses challenges in textual analysis and offers insights into author identification.

2. Objectives

- Develop a robust system to predict the author of a text passage.
- Ensure accuracy and reliability in predictions.
- Provide an easy-to-use interface for end-users.
- Deploy the system for scalable and practical use.

3. Scope

- **Inclusions:** Feature-based author prediction using supervised learning, user interface for input, and deployment on a cloud-hosted environment.

- **Exclusions:** Real-time text processing, feature engineering, and dataset creation.

4. Methodology

4.1 Data Collection and Preprocessing

- Analyze and understand the dataset, ensuring it is clean and structured for use.
- Utilize predefined features without additional text processing.

4.2 Machine Learning Modeling

- Train models such as Random Forest and Gradient Boosting on the dataset.
- Perform hyperparameter tuning and validate the model using cross-validation.

4.3 User Interaction

- Develop a simple user interface (CLI or desktop-based) for input and prediction display.

4.4 Deployment

- Host the application on Render or similar server environments without containerization.

5. Technology Stack

- **Programming Languages:** Python (Backend)
- **Frameworks and Libraries:** scikit-learn for machine learning.
- **Hosting Platform:** Render for deployment.

6. Results and Evaluation

- Evaluate the model performance using metrics such as accuracy, precision, recall, and F1 score.
- Summarize the key findings and validate the effectiveness of the system.

7. Conclusion

The Author Prediction System for the Avila Bible achieves its objective of predicting authorship with high accuracy. It provides an efficient and user-friendly solution for text analysis and attribution.

8. Future Work

- Integrate additional features for more granular predictions.
- Explore advanced machine learning models such as deep learning techniques.
- Expand the system to support real-time text processing and predictions.

9. References

1. Dataset Documentation
2. scikit-learn Documentation
3. Render Deployment Guides

4. https://r.search.yahoo.com/_ylt=AwrKAjDiFnlnLgIAL8u7HAX.;_ylu=Y29sbwNzZzMEcG9zAzEEdnRpZAMEc2VjA3Ny/RV=2/RE=1737198563/RO=10/RU=https%3a%2f%2fwww.geeksforgeeks.org%2fregression-using-catboost%2f/RK=2/RS=3.a4WMVN6NSFWgxSqzZpt1br6xc-
5. https://catboost.ai/docs/en/concepts/python-reference_catboostregressor