

Project Proposal

Forecasting Student PET English Test Achievement Using Placement and Diagnostic Test Data

Fatimah Nowh Aljaroudi

Master's Student, Master's Student, Department of Artificial Intelligence

Midocean University

1. Project Overview

This project aims to develop a predictive model to forecast student performance on the PET English test. By integrating data from placement tests, diagnostic assessments, and records of English course participation, the study addresses the need for early identification of students at risk of underperforming. Accurate predictions will enable educators to design timely interventions and personalized learning strategies, ultimately improving language proficiency outcomes.

2. Project Goal

- Develop a machine learning model that accurately forecasts PET English test scores.
- Improve early identification of students needing additional support by analyzing placement test, diagnostic test, and course participation data.
- Provide actionable insights that can inform curriculum improvements and targeted interventions.

3. Approach

- **Data Collection:**
Gather historical data from placement tests, diagnostic assessments, and detailed logs of English course participation (e.g., attendance records, grades, course types). Data sources may include institutional records and digital learning platforms.
- **Data Processing:**
Clean and preprocess the collected data by handling missing values, normalizing scores, and performing feature engineering to extract relevant predictors.

- **Model/Method Selection:**
Evaluate several machine learning approaches including regression models, decision trees, and neural networks. Selection will be based on model performance and interpretability.
- **Testing and Evaluation:**
Assess the model using evaluation metrics such as RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), and R^2 . Cross-validation techniques will be employed to ensure robust performance and generalizability.

4. Project Steps and Timeline

- literature review (Week 3-4)
- Data Collection and Preparation (Week 5-6)
- Model Building and Training (Week 7-8)
- Testing and Evaluation (Week 9)
- Final Adjustments and Reporting (Week 10)

5. Expected Results

- A robust predictive model that accurately forecasts PET English test scores.
- Identification of key factors that influence test performance, offering insights into student learning behavior.
- A framework that can be integrated into educational systems for early intervention and personalized support, ultimately enhancing student outcomes.

6. Resources Needed

- **Software and Tools:** Python programming environment, libraries such as scikit-learn and TensorFlow, and data visualization tools (e.g., Matplotlib, Seaborn).
- **Datasets:** Institutional records of placement tests, diagnostic assessments, and English course participation data.
- **Support:** Access to computing resources for model training, guidance from academic supervisors, and potential collaboration with educational institutions for data access.

7. References

Recent studies on educational data mining and predictive analytics in language testing.

Key texts on machine learning applications in education.

Articles and conference papers on the use of diagnostic testing and course data to forecast academic performance.