**Employee Engagement and Learning Behavior**

Contents

[1 Problem Statement: 2](#_Toc179812709)

[2 Overview: 2](#_Toc179812710)

[2.1 Objectives: 2](#_Toc179812711)

[3 Architectural Diagram: 3](#_Toc179812712)

[4 Solutions: 4](#_Toc179812713)

[4.1 MERN Stack Implementation 4](#_Toc179812714)

[4.1.1 Admin Login: 4](#_Toc179812715)

[4.1.2 Employee Login: 4](#_Toc179812716)

[4.2 Data Engineering 7](#_Toc179812717)

[4.2.1 Data Preparation: 7](#_Toc179812718)

[4.2.2 Reporting: 7](#_Toc179812719)

[4.2.3 Final Reporting Table: 7](#_Toc179812720)

[4.3 Data Science 8](#_Toc179812721)

[4.3.1 Target Column Preparation: 8](#_Toc179812722)

[4.3.2 Analysis: 8](#_Toc179812723)

[4.3.3 Model Development: 11](#_Toc179812724)

[5 Conclusion 11](#_Toc179812725)

# Problem Statement:

Create a system that tracks employee engagement with learning materials, focusing on metrics such as time spent on the platform, and quiz scores.

# Overview:

The project involves building a comprehensive dashboard for tracking employee engagement and learning behaviours. The system will provide both employee and admin interfaces, allowing for detailed analysis of engagement levels across teams and department

## Objectives:

* Track time spent on learning materials, and quiz scores.
* Enable employees to provide feedback on courses and learning materials.
* Design a robust database to store engagement data and feedback.

# Architectural Diagram:

A diagram of a diagram

Description automatically generated

# Solutions:

## MERN Stack Implementation

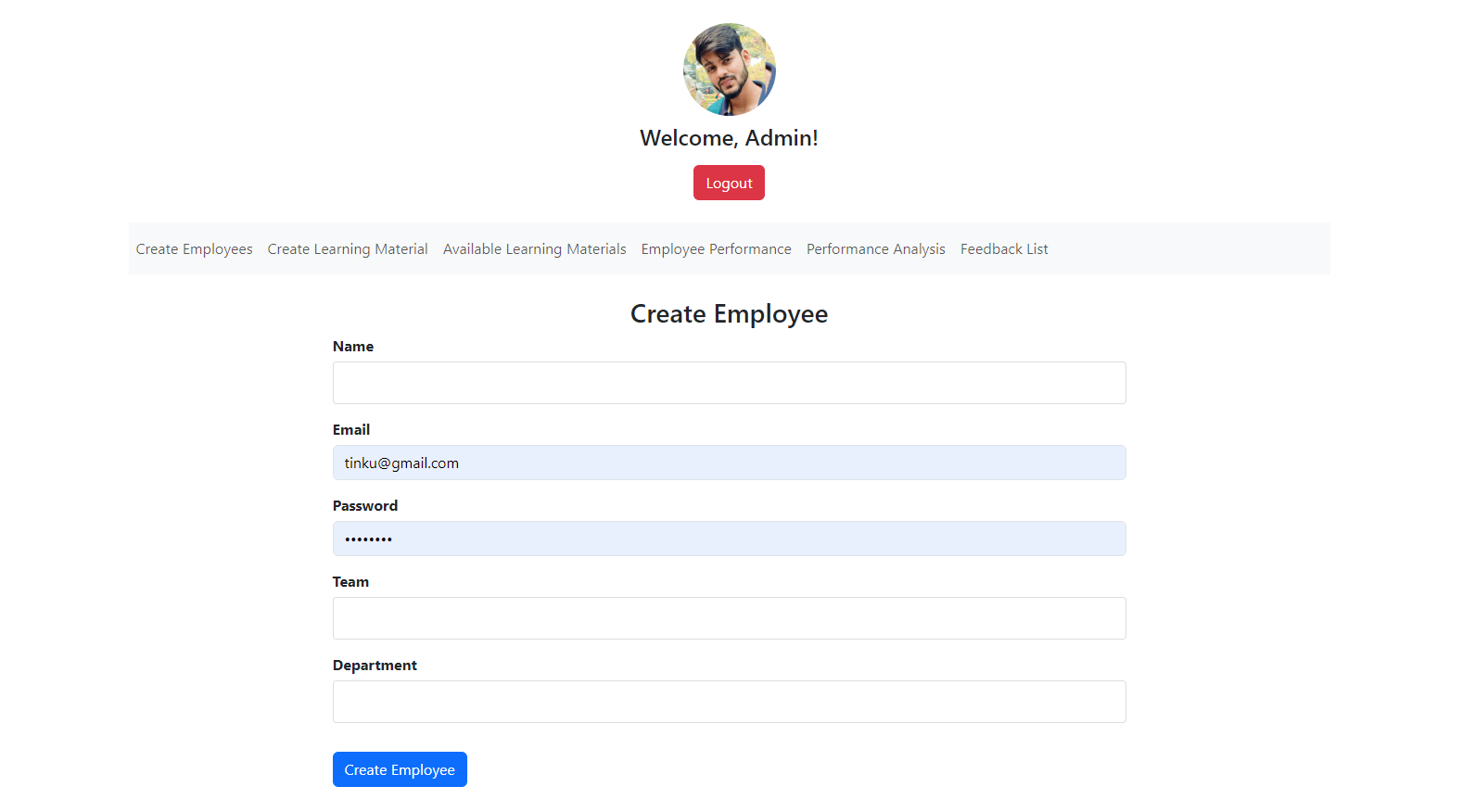
The project utilizes the MERN (MongoDB, Express.js, React.js, Node.js) stack to create the following features:

### Admin Login:

* Create and edit employee profiles.
* Create courses and quizzes associated with those courses.
* Analyse performance through visualizations of top 5 employees by score and time spent, top 5 departments by average score, and top 5 courses by average score.
* View employee performance metrics such as scores and time spent.
* Access and review feedback submitted by employees.

### Employee Login:

* View available courses and access learning materials.
* Take quizzes for enrolled courses and mark them as completed upon submission.
* Monitor personal performance metrics for completed courses.
* Submit feedback for each course attended.



Admin Page

A screenshot of a computer

Description automatically generated

Employee page

A screenshot of a feedback form

Description automatically generated

Feedback Form

*A screenshot of a computer

Description automatically generated*

Employee Performance Analysis

## Data Engineering

In this phase, we connected to our MongoDB database and generated synthetic data using the Faker library, creating 25,000 entries to simulate engagement metrics.

### Data Preparation:

* Extracted and saved raw data from the MongoDB database.
* Cleaned the data by checking for null values, adjusting data types, and dropping unnecessary columns.
* Constructed fact tables (Scores, Feedback) and dimensional tables (Users, Quizzes, Courses).

### Reporting:

#### Created summary reports:

* **Feedback Summary Report:** Count of feedbacks for each title.
* **User Feedback Summary Report**: Count of feedbacks given by users.
* **User Engagement Report:** Metrics for each employee, including total quizzes taken, total score, average score, and total time spent.

#### Calculated an engagement score based on average ratings, total quizzes taken, average scores, total time spent, and total feedback count.

#### Provided insights on:

* Best and worst performing departments based on scores.
* Top and least effective courses based on time spent and scores.
* Top-performing employees based on scores and time spent.

### Final Reporting Table:

#### The data prepared for the Data Science phase was stored in final\_reporting\_table.csv, containing the following columns:

* + Rating
  + Name Engagement
  + Department
  + Total Quizzes Taken
  + Total Score
  + Average Score
  + Total Time Spent
  + Engagement Score

## Data Science

In the Data Science phase, we aimed to predict future employee engagement levels based on historical behaviour patterns.

### Target Column Preparation:

* Bucketing the [engagement score] column into categories: low, medium, and high engagement.
* Label-encoded these categories as [engagement level numeric], our target column.

### Analysis:

#### Analysed the distribution of the target column.

#### Conducted hypothesis testing to determine the relationship between various metrics and engagement levels:

* + Higher average ratings correlate with higher engagement.
  + Increased feedback count correlates with higher engagement.
  + More quizzes taken correlate with higher engagement.
  + Increased time spent correlates with higher engagement.
  + Higher average scores correlate with higher engagement.

A graph of a distribution of engagement levels

Description automatically generated

Distribution of target column

*A diagram of a graph

Description automatically generated*

*Average score by engagement level*

*A graph of blue rectangular shapes

Description automatically generated with medium confidence*

*Average Total time Spent By Engagement level*

*A graph of a number of people

Description automatically generated*

*Count of employees by engagement level and department*

*A graph of a graph

Description automatically generated with medium confidence*

*Total feedback count by engagement level*

### Model Development:

* + Separated features from the target variable and split the data into training and test sets.
  + Standardized numerical features for improved model performance.
  + Applied logistic regression, initially achieving 100% accuracy, indicating overfitting.
  + Utilized a Random Forest Classifier, resulting in training accuracy close to 1 and test accuracy at 93%, also indicating overfitting.
  + Reduced model complexity by adjusting [estimators] and limiting [max depth].
  + Achieved a final model accuracy of 87% for training and 84% for the test set.
  + Visualized model performance using confusion matrices and AUC-ROC curves for each class (high, medium, low).

# Conclusion

The "Employee Engagement and Learning Behaviour" project effectively combines MERN stack development, data engineering, and data science techniques to provide insights into employee learning engagement. By leveraging these technologies, the project aims to create personalized engagement strategies for HR departments based on predictive analytics.

**THE END**