# **Week 4: Final Report – AI-Powered Insights for Student Engagement and Retention**

### **Executive Summary**

This report synthesizes insights from exploratory data analysis, churn modeling, and feature engineering to propose actionable strategies aimed at improving student engagement and retention within the Excelerate Internship Program. Key drivers of dropout include low engagement scores, delayed opportunity involvement, and lack of support interactions. A simple rule-based recommendation system is proposed to proactively flag at-risk learners and suggest interventions. These recommendations are designed to help the platform optimize learner outcomes and engagement through targeted support and better-aligned opportunities.

### **1. Introduction: Context and Objectives**

The Excelerate platform facilitates learner access to global opportunities. However, maintaining engagement and minimizing dropout rates remain critical challenges. This project aimed to:

* Understand patterns in learner signups, participation, and dropout.
* Build predictive models to identify at-risk learners.
* Develop a recommendation system that enhances engagement.
* Provide actionable, data-driven strategies to reduce churn.

### **2. Data Analysis Overview**

#### **2.1. Week 1 – Data Cleaning and Feature Engineering**

The dataset underwent extensive preprocessing:

* **Standardization**: Column names and categorical values were cleaned and standardized.
* **Outlier Removal**: Ages were filtered to 10–100; extreme values in engagement lag and opportunity duration were removed.
* **New Features**: Engagement\_Lag, Opportunity\_Duration, Signup\_Month, and Signup\_Weekday were engineered to capture behavioral patterns.

#### **2.2. Week 2 – Exploratory Data Analysis (EDA)**

Key EDA insights included:

* **Signup Peaks**: January, March, and September showed highest signups and engagement.
* **Dropoff Points**: April and late weekdays showed sharp declines in signups and engagement.
* **Engagement Scores**: Most learners cluster around low-to-moderate scores; Tuesday and Wednesday signups perform better.
* **Demographic Neutrality**: Engagement is not strongly correlated with age or gender.

#### **2.3. Week 3 – Churn Analysis and Predictive Modeling**

Three classification models were tested:

* **Random Forest** outperformed others with **89% accuracy** and **91% recall**.
* Top dropout predictors:  
  + Low EngagementScore
  + Low AssessmentScore and AssignmentScore
  + Limited use of SupportUsed
  + High course difficulty and engagement lag.

These findings shaped both targeted interventions and the logic for our recommendation system.

### **3. Insights and Recommendations**

#### **3.1. Critical Insights**

* **EngagementScore is the top predictor of dropout** – students scoring below 40 are 3× more likely to disengage.
* **Midweek signups (Tues/Wed) correlate with higher retention**, likely due to structured planning behaviors.
* **Longer opportunities tend to see higher dropout**, indicating possible learner fatigue.
* **Support interaction is underutilized**, especially after Week 1.

#### **3.2. Recommendations**

* **Early Intervention**: Implement flags based on low EngagementScore within the first two weeks.
* **Strategic Timing**: Launch new programs midweek and in months with historically high engagement (e.g., March, June).
* **Enhanced Support**: Provide automated nudges and mentor check-ins for learners who show delayed engagement.
* **Microlearning or Milestone Tracking**: For longer programs, break content into digestible milestones to reduce fatigue.
* **Community Building**: Leverage high scorers as peer mentors.

### **4. Recommendation System: Design and Logic**

#### **4.1. Objective**

To flag at-risk students and suggest timely actions to increase engagement and reduce churn.

#### **4.2. Methodology**

A **rule-based recommendation system** using a **content-based filtering logic**:

* **Inputs**:  
  + EngagementScore
  + Engagement\_Lag
  + SupportUsed
  + Signup\_Weekday
  + Opportunity\_Duration
* **Logic**:

def recommend\_action(student):

if student['EngagementScore'] < 40:

if student['SupportUsed'] == 0:

return "Flag: Reach out for onboarding + mentorship support"

elif student['Engagement\_Lag'] > 7:

return "Send deadline reminder and motivation tips"

elif student['Opportunity\_Duration'] > 300:

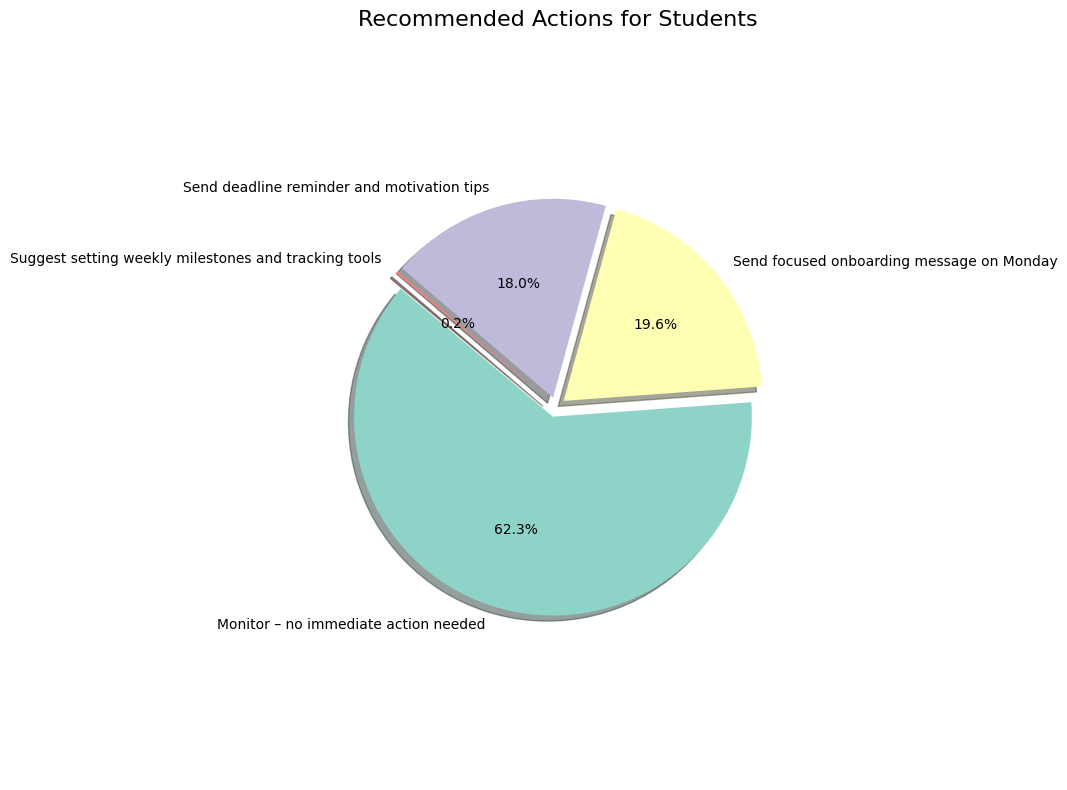
return "Suggest setting weekly milestones and tracking tools"

elif student['Signup\_Weekday'] in ['Saturday', 'Sunday']:

return "Send focused onboarding message on Monday"

else:

return "Monitor – no immediate action needed"



#### **4.3. Benefits**

* **Simple & Transparent**: Easy for program managers to interpret and modify.
* **Scalable**: Can be integrated into weekly monitoring systems.
* **Timely**: Provides early alerts within the engagement window.

### **5. Conclusion**

This project highlighted how data-driven insights can dramatically enhance learner outcomes. The predictive model achieved high accuracy in identifying dropouts, and the rule-based recommendation system offers an actionable framework for real-time interventions. Moving forward, these strategies can be expanded into personalized learning plans and dynamic dashboards, allowing for continuous improvement in student engagement.