# Big Data Analytics

## Exam Anxiety Prediction Model

Understanding Student Stress Before Exams

Capstone Project | Industry: Education | December 2025

# The Problem

📊 **45-60% of students** experience significant exam anxiety

😴 **Sleep deprivation** becomes severe 1 week before exams

🔴 **Early detection** can enable targeted interventions

📱 **Big Data** from multiple sources can identify patterns

# Project Objectives

- **Objective 1** Predict high-stress students before exams
- **Objective 2** Track study time vs stress correlation
- **Objective 3** Analyze social media mood trends
- **Objective 4** Detect sleep pattern changes
- **Objective 5** Suggest personalized interventions

# Dataset Characteristics

### 📊 Sample Size

**100 students**

### 📋 Variables

**27 features**

### 🎯 Target

**High Risk Classification**

Data collected across demographics, sleep patterns, study habits, social media behavior, stress indicators, and mental health metrics

# Data Features Used

### Behavioral Metrics

- Sleep disruption
- Study hours/patterns
- Social media usage
- Mood tracking

### Physiological Indicators

- Anxiety levels
- Cortisol levels
- Physical symptoms
- Stress scores

# Machine Learning Approach

1. **Data Preprocessing:** Cleaning, scaling, feature engineering
2. **Model Selection:** 3 classifiers tested
3. **Training:** 70-30 train-test split with cross-validation
4. **Evaluation:** Accuracy, ROC-AUC, classification metrics
5. **Interpretation:** Feature importance & insights analysis

# Model Performance Results

| Logistic Regression | Random Forest | Gradient Boosting |
| --- | --- | --- |
| **89.66%** | **96.55%** | **94.83%** |
| Accuracy | Accuracy | Accuracy |

🏆 **Best Model: Random Forest with 96.55% accuracy**

# Top Predictive Features

1st Baseline Anxiety (Most Important)

2nd Stress Management Score

3rd Sleep Disruption Severity

4th Physical Symptoms Count

5th Social Support Rating

These 5 features account for 62% of prediction variance

# 🎯 Use Case 1: Individual Prediction

**Predict high-stress students** before exams occur

For each student, the model provides:

- Risk classification (High/Low)
- Risk probability (0-100%)
- Key stress indicators
- Personalized intervention suggestions

# 📊 Use Case 2: Batch Analysis

**Process entire cohorts** to identify patterns

Analysis capability:

- Identify proportion of high-risk students
- Compare risk by department/major
- Track trends across exam periods
- Allocate resources effectively

# 💡 Use Case 3: Interventions

**Personalized stress management** recommendations

Recommendations based on indicators:

- Sleep recovery programs
- Meditation & mindfulness coaching
- Exercise & wellness plans
- Counseling & social support

# Key Data Insights

🛌 Students losing >2 hours sleep: **3.2x higher risk**

🧘 Meditation practice: **Reduces risk by 45%**

👥 Strong social support: **Protective factor for 68% of students**

📚 Study hours >35/week: **Increases stress 2.1x**

# Business Impact & Value

- ✓ **Early Detection:** Identify at-risk students 1-2 weeks before crisis
- ✓ **Targeted Support:** Direct counseling resources to highest-need students
- ✓ **Improved Outcomes:** Reduce anxiety-related dropouts and academic failure
- ✓ **Scalability:** Apply model across entire institution
- ✓ **Data-Driven Decision Making:** Replace reactive with proactive counseling

# Thank You

Questions?

Capstone Project Submission | Big Data Analytics

December 2025