Data Science Capstone Project: Overstimulation Behaviour and Lifestyle Analysis

Project Overview

Overstimulation has become a growing concern in today's digital world, where individuals are constantly exposed to excessive stimuli from social media, digital devices, and environmental factors. Understanding how overstimulation affects human behavior, mental health, and lifestyle can help in developing better intervention strategies. In this capstone project, students will analyze the **Overstimulation Behavior and Lifestyle Dataset** to uncover insights, trends, and patterns related to overstimulation and its impact on daily life.

Dataset Details

Dataset Name: Overstimulation Behavior and Lifestyle Dataset
Source: Kaggle (https://www.kaggle.com/datasets/miadul/overstimulation-behavior-and-lifestyle-dataset)
Number of Records: Varies (Check dataset)
Number of Features: 20

Key Features

- **Age, Gender** – Demographic information of respondents

- **Screen Time (hours/day)** – Number of hours spent on digital screens

- **Social Media Usage (hours/day)** – Time spent on various social media platforms

- **Multitasking Habit (Yes/No)** – Whether the individual frequently multitasks

- **Sleep Duration (hours/day)** – Average sleep duration per day

- **Physical Activity Level (Low/Moderate/High)** – Self-reported activity levels

- **Mental Health Rating (1-10)** – Self-reported mental health status

- **Cognitive Fatigue Level (1-10)** – Self-reported cognitive fatigue score

- **Work/Study Efficiency (%)** – Productivity levels while working or studying

- **Stress Level (1-10)** – Self-reported stress levels

- **Diet Type (Healthy/Moderate/Unhealthy)** – Type of diet followed

- **Hobbies (Yes/No)** – Whether the person engages in hobbies regularly

- **Family/Social Interaction Level (1-10)** – Quality of personal interactions

- **Overstimulation Score (1-10)** – Overall overstimulation self-rating

Scenario

A wellness research institute is studying the impact of overstimulation on human lifestyle and mental well-being. The institute aims to identify key factors contributing to overstimulation and provide recommendations for healthier digital habits. Your role as a data scientist is to analyze this dataset, identify patterns, and provide data-driven insights and recommendations.

Problem Statements

1. How does screen time correlate with cognitive fatigue and stress levels?

2. What impact does social media usage have on mental health and work efficiency?

3. Does physical activity play a role in reducing overstimulation effects?

4. Are there significant differences in overstimulation scores based on demographic factors (age, gender)?

5. How do sleep duration and multitasking habits contribute to cognitive fatigue?

6. What lifestyle factors help in minimizing overstimulation?

Project Deliverables

1. Data Exploration & Cleaning

- Load and clean the dataset (handle missing values, outliers, inconsistencies)

- Perform exploratory data analysis (EDA) to understand data distributions

- Visualize key trends using graphs and plots

2. Statistical & Correlation Analysis

- Identify correlations between overstimulation score and other variables

- Conduct hypothesis testing to validate key relationships

3. Machine Learning Model

- Train predictive models to estimate overstimulation scores based on lifestyle factors

- Compare different models (e.g., Linear Regression, Decision Trees, Random Forest)

4. Insights & Recommendations

- Summarize key findings with actionable insights

- Propose recommendations for healthier lifestyle habits

5. Final Report & Presentation

- Document the entire process, methodology, and findings

- Create a compelling presentation with visualizations

- Provide a brief video or live presentation explaining key insights

---

Evaluation Criteria

Criteria	Weightage
Data Cleaning & Preprocessing	20%
Exploratory Data Analysis (EDA)	20%
Statistical Analysis & Insights	20%
Machine Learning Model	15%
Report & Presentation	25%

Tools & Technologies

- **Python**: Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn

- **Jupyter Notebook / Google Colab**

- **Power BI / Tableau (Optional for visualization)**

Submission Guidelines

- Submit a well-documented Jupyter Notebook/ Pycharm with the full analysis.

- Provide a PDF report summarizing findings, visualizations, and recommendations. (business Memo).

- Prepare a 5–10-minute presentation (live or recorded) to be uploaded on LinkedIn as an Article.

- Paste the LinkedIn Post link in the presentation at the end before submitting.

LinkedIn Project Posts for Reference
1. https://www.linkedin.com/posts/namratadutta03_stockanalysis-oesoninternship-dataanalysis-activity-7175003464878428160-Kkt2?utm_source=share&utm_medium=member_desktop&rcm=ACoAABrueEsBWe7Fe6k69yDIqFh-gvCSjkPwBvU
2. https://www.linkedin.com/posts/mike-barbiere-7989b51a0_dataanalysis-datavisualization-python-activity-7166874519750586368-nMnE?utm_source=share&utm_medium=member_desktop&rcm=ACoAABrueEsBWe7Fe6k69yDIqFh-gvCSjkPwBvU

LinkedIn Articles Posts for Reference

1. https://www.linkedin.com/pulse/become-data-engineer-2024-dulmi-sapna-sehani--zcilc/?trackingId=8EUGGPnVS2uFTfW7L1ptZw%3D%3D

2. https://www.linkedin.com/pulse/man-behind-computer-journey-through-cybersecurity-oeson-a-mewborn-l5p8e

3. https://www.linkedin.com/pulse/my-internship-experience-oeson-atakan-erdogan-tjbtf

Testimonial Video

https://www.linkedin.com/posts/oesonlearning_datascience-coursereview-studenttestimonial-activity-7307226678139719680-4QnM?utm_source=share&utm_medium=member_desktop&rcm=ACoAABrueEsBWe7Fe6k69yDIqFh-gvCSjkPwBvU