STUDENT ATTENTIVENESS ANALYSIS IN VIRTUAL CLASSROOMS USING DISTRACTION DETECTION

Submitted in partial fulfilment for the course DEEP LEARNING AND MLOPS

Submitted By

Team No: 29

Amrutha K V: 243107, MSc Data Analytics with Computational Science.

Anakha V Wills: 243303, M.Sc. Data Analytics with Geoinformatics.

Dhanush B. R : 243307, M.Sc. Data Analytics with Geoinformatics.

To

Dr. Manoj Kumar T. K.

School of Digital Sciences

Digital University Kerala

Aim

The primary aim of this project is to detect student fatigue and distraction in virtual classrooms by analyzing attentiveness through facial cues and head movements. In a remote learning environment, instructors cannot observe students' body language, making it difficult to assess engagement levels. This system addresses that challenge by classifying students as "Focused" or "Not Focused" based on head pose estimation using Euler angles such as pitch, roll, and yaw. The goal is to enhance learning outcomes by providing real-time feedback and post-class analysis, supporting instructors, students, and institutions.

Methodology and Results

The methodology involves collecting webcam videos from participants, we are using smoothed and facially cropped video version of the University of Texas at Arlington Real-Life Drowsiness Dataset (UTA-RLDD), tailored for use in temporal deep learning models. It contains recordings from 60 subjects across two classes: non-drowsy and drowsy. Videos are provided at a frame rate of 10 frames per second (fps) and a fixed resolution of 224×224 pixels, ensuring consistency across samples. The dataset includes video clips of varying lengths — 5, 10, 20, 30, and 60 seconds — making it versatile for training models that analyze temporal patterns in facial behaviour related to alertness and fatigue. Facial landmarks are extracted using Dlib, and head pose features are calculated. These features are used to train an LSTM model with four hidden layers and approximately 145,000 parameters. The model captures temporal patterns in head movement to predict focus levels. Real-time predictions are refined using frame-by-frame analysis and moving average smoothing. We expect an accuracy of 90.2%, showing strong potential for integration into virtual learning platforms, online exams, and corporate training setups.