



# CSE523 MACHINE LEARNING

**Weekly Report 3: 17 February 2024**

**WDL**

Name	Enrollment No.
Paridhi Jain	AU2120226
Nevil Jobanputra	AU2140209
Saahil Doshi	AU2140106
Jash Parikh	AU2140108

**Project 8:** Athlete performance in collegiate basketball: Predicting match line-up (Visualization on Dashboard)

## **Week 3: Data Collection and Review**

### **Accomplishments:**

- Collected relevant data sources, including athlete performance metrics, training details, sleep patterns, and emotional-mental state information from the author of the research paper of reference.
- Reviewed the quality and completeness of the collected data, identifying any inconsistencies or missing values.
- Explored potential data augmentation techniques to balance the dataset and address class imbalances.
- Prepared presentation of the progress discussing approach and datasets.

### **Challenges:**

- Ensuring the completeness and accuracy of the collected data from different sources.
- To address data preprocessing challenges, such as handling missing values and outliers, which were found in huge numbers in the dataset.
- Understanding the nuances and interrelationships between different variables in the dataset.

### **Next Steps:**

- Implement robust data preprocessing techniques to handle missing values and outliers effectively.
- Explore strategies for imputing missing data, such as mean imputation, interpolation, or advanced imputation methods like K-nearest neighbors.
- Continue to assess the balance of the dataset and investigate methods like oversampling or undersampling to address class imbalances.