Proposal: Chess Blunder Analysis Using Multivariate and Linear Algebraic Approaches

Author: Vandit Goel
Student ID: 1002245699
Email: vxg5699@mavs.uta.edu

This project investigates the phenomenon of blunders in chess by analyzing when and why players make them, with a focus on combining multiple explanatory factors by analyzing large-scale game data from the publicly available **Lichess database** (90,000,000 games/month). Each move will be evaluated using the **Stockfish engine** to quantify **shifts in evaluation and win/draw/loss probabilities**, which will serve as a baseline for identifying blunders. Alongside these engine-derived features, additional factors such as the sharpness of the position, the type of piece moved, and temporal factors such as time remaining and time spent per move will be incorporated into the analysis. By combining these dimensions, the study aims to capture the complex interplay between position difficulty, human decision-making under pressure, and move selection errors. A further goal is to develop a **predictive model, using a neural network**, that estimates the likelihood of a blunder occurring in a given position and its context.

The project will employ a range of mathematical and computational methods to represent, analyze, and model the data. Board positions and their associated features will be **encoded as matrices or tensors**, allowing for structured analysis of the relationships among variables. Optimization methods and matrix factorizations will be used in model training and in reducing the dimensionality of the feature space, ensuring more efficient computation while preserving the most important information. Clustering methods built on kernel matrices will be applied to group similar blunder types, such as tactical oversights versus strategic misjudgments. Finally, supervised learning and neural network techniques will be used to train, validate, and test a predictive system that links position characteristics to blunder probabilities. The project is therefore `experimental` in nature: it will involve extracting and preprocessing a large dataset, applying mathematical and computational tools to analyze patterns of mistakes, and evaluating the effectiveness of predictive modeling for understanding and forecasting blunders in chess.