COMP721 – Machine Learning

Project – The National Basketball Association (NBA) Prediction

Due Date: 14/11/2025
- Submit a typed assignment on Moodle
- Submit in **PDF format** not **Zipped files**
- The project can be done in groups of 3 or 5 students

This course project is an opportunity for you to explore a Machine Learning problem using a real-world dataset. The idea is to apply multiple Machine Learning techniques or major variations of a single approach (typically, different groups will explore different approaches), to compare the results, and to summarize your findings in a report.

Given a dataset (***databasebasketball.zip***) which contains 2004-2005 NBA and ABA statistics for:

- Player regular season stats
- Player regular season career totals
- Player playoff stats
- Player playoff career totals
- Player all-star game stats
- Team regular season stats
- Complete draft history
- coaches_season.txt - NBA coaching records by season
- coaches_career.txt - NBA career coaching records

In the project, you are required to model Machine Learning techniques to:

- find out who the outstanding players, outlier detection on the players.
- predict the game outcome, given two teams.

The following submissions are required:

- Project report about **6 pages** using the following format (Introduction, Methods and Techniques, Results and Discussion, Conclusion, References).
- GitHub link for the project code.

For more information: https://www.basketball-reference.com/players/j/jamesle01.html