Tennis match forecasting; how AI can beat book makers on their own platforms

Patrick Schall* and Vahid Toomani†

(Dated: February 28, 2024)

* patrick.schall@gmail.com

† vahidtoomani2002@gmail.com

CONTENTS

I. Introduction 3

A. International tournaments and book makers 3

B. Elo system 3

C. objectives 3

II. Exploratory data analysis (EDA) 4

A. Data fetch 4

B. Data enrichment & cleaning 4

C. Data statistics and visualisations 4

D. Strategies 4

E. Feature selection 5

III. ML models and learning 5

A. Model selection 5

B. Random forest 5

C. Ada boost 5

D. Support vector Mmachine 5

E. Voting 5

F. Dense neural network 5

IV. Post learning analysis & packaging 5

A. Uncertainty cutoff 5

B. Cutoff optimization 5

V. Conclusion 5

I. INTRODUCTION

Intro:

- Context of the project's integration into your business.
- From a technical point of view.
- From an economic point of view.
- From a scientific point of view.

A. International tournaments and book makers

ATP, Pinnacle sports, Bet 365

B. Elo system

Elo rate

C. objectives

Goals

- What are the main objectives to be achieved? Describe in a few lines.
- For each member of the group, specify the level of expertise around the problem addressed?
- Have you contacted business experts to refine the problem and the underlying models? If yes, detail the contribution of these interactions.
- (Are you aware of a similar project within your company, or in your entourage? What is its progress? How has it helped you in the realization of your project? How does your project contribute to improving it?).

II. EXPLORATORY DATA ANALYSIS (EDA)

A. Data fetch

source

B. Data enrichment & cleaning

Which parts of data dropped or added?

C. Data statistics and visualisations

Put graphs and statistics of the data here.

D. Strategies

Strategies

E. Feature selection

III. ML MODELS AND LEARNING

A. Model selection

B. Random forest

C. Ada boost

D. Support vector Mmachine

E. Voting

F. Dense neural network

IV. POST LEARNING ANALYSIS & PACKAGING

A. Uncertainty cutoff

B. Cutoff optimization

V. CONCLUSION