Independent Verification and Validation Report

General Information

State	Project Name	Program Name	Progress Report Date	POC Name
Virginia	TestProject	E&E	12-15-15	Bob

POC Email	Submitter Name	Submitter Role	Submitter Email	Submitter Phone
Bob@va.gov	Jim	Admin	Jim@va.gov	7035721234

Activity 1 Consult Date	RFP Release Date	IV&V Onboard Date	Next Progress Report Date
04-12-12	12-15-12	12-15-12	12-15-12

Executive Summary

Extensive research has been conducted into the modelling of professional tennis matches. Most current approaches take advantage of the hierarchical structure of the tennis scoring system to define stochastic models, based on Markov chains. These models use only the probability of each of the players winning a point on their serve to compute their respective probabilities of winning the match. Consequently, a variety of factors that contribute to the outcome of a match are ignored. We propose a supervised machine learning approach that uses historical player performance across a wide variety of statistics to predict match outcomes. We define a novel method of extracting 22 features from raw historical data, including abstract features, such as player fatigue and injury. Using the resulting dataset, we develop and optimise models based on two machine learning algorithms: logistic regression and artificial neural networks. When evaluated on a test set of 6315 ATP matches played in the years 2013-2014, our models outperform Knottenbelt's Common-Opponent model, the current state-of-the-art in stochastic modelling. Our neural network generates a return on investment of 4.35% when in competition with the betting market, an improvement of about 75%. We believe that the use of machine learning will lead to innovation in the field of tennis modelling.

Project Management Office Status

Total Budget	Earned Value(EV)	Budget Variance(%)	Schedule Variance(%)	Other
5%	15%	25%	35%	10%

Life Cycle Status and Schedule

Care Management (Status: Pre-R1)

Target App. Date	Target Dev. Start	Target R1	Target R2	Target Go Live	Target R3
12-22-12	12-22-12	12-22-12	12-22-12	12-22-12	12-22-12

Contractor Management (Status: R1)

Target App. Date	Target Dev. Start	Target R1	Target R2	Target Go Live	Target R3
12-22-12	12-22-12	12-22-12	12-22-12	12-22-12	12-22-12

Third Party Liability (Status: R2)

Target App. Date	Target Dev. Start	Target R1	Target R2	Target Go Live	Target R3
12-22-12	12-22-12	12-22-12	12-22-12	12-22-12	12-22-12

FFS Claims and Adjudication (Status: R3)

Target App. Date	Target Dev. Start	Target R1	Target R2	Target Go Live	Target R3
12-22-12	12-22-12	12-22-12	12-22-12	12-22-12	12-22-12

Risks

Risk 5 (ID: 5)

Description: Extensive research has been conducted into the modelling of professional tennis matches. Most current approaches take advantage of the hierarchical structure of the tennis scoring system to define stochastic models, based on Markov chains. These models use only the probability of each of the players winning a point on their serve to compute their respective probabilities of winning the match. Consequently, a variety of factors that contribute to the outcome of a match are ignored. We propose a supervised machine learning approach that uses historical player performance across a wide variety of statistics to predict match outcomes. We define a novel method of extracting 22 features from raw historical data, including abstract features, such as player fatigue and injury. Using the resulting dataset, we develop and optimise models based on two machine learning algorithms: logistic regression and artificial neural networks. When evaluated on a test set of 6315 ATP matches played in the years 2013-2014, our models outperform Knottenbelt’s Common-Opponent model, the current state-of-the-art in stochastic modelling. Our neural network generates a return on investment of 4.35% when in competition with the betting market, an improvement of about 75%. We believe that the use of machine learning will lead to innovation in the field of tennis modelling.

Probability	Impact	Risk Score	Target Resolution Date	Status
2	2	1	11-15-12	Good

Risk 2 (ID: 2)

Description: Hey everyone! I publish a few blog posts each week on JavaScript/Web Development, and I wanted to share my articles from June. They focus on JavaScript and Node.js, but there is also stuff on Express, React Native, CSS Flexbox, and Pug. I try to make tutorials that are interesting, and that result in you building something cool - so hopefully this is helpful.

Probability	Impact	Risk Score	Target Resolution Date	Status
2	12	1	11-12-12	Bad

Recommendations

Recommendation #: 4 (Date of Recommendation: 05-06-12, Resolved?: No)

Recommendation: Hey everyone! I publish a few blog posts each week on JavaScript/Web Development, and I wanted to share my articles from June. They focus on JavaScript and Node.js, but there is also stuff on Express, React Native, CSS Flexbox, and Pug. I try to make tutorials that are interesting, and that result in you building something cool - so hopefully this is helpful.

Comments: There are no comments right now
Recommendation #: 4 (Date of Recommendation: 10-04-01, Resolved?: Yes)

Recommendation: Extensive research has been conducted into the modelling of professional tennis matches. Most current approaches take advantage of the hierarchical structure of the tennis scoring system to define stochastic models, based on Markov chains. These models use only the probability of each of the players winning a point on their serve to compute their respective probabilities of winning the match. Consequently, a variety of factors that contribute to the outcome of a match are ignored. We propose a supervised machine learning approach that uses historical player performance across a wide variety of statistics to predict match outcomes. We define a novel method of extracting 22 features from raw historical data, including abstract features, such as player fatigue and injury. Using the resulting dataset, we develop and optimise models based on two machine learning algorithms: logistic regression and artificial neural networks. When evaluated on a test set of 6315 ATP matches played in the years 2013-2014, our models outperform Knottenbelt's Common-Opponent model, the current state-of-the-art in stochastic modelling. Our neural network generates a return on investment of 4.35% when in competition with the betting market, an improvement of about 75%. We believe that the use of machine learning will lead to innovation in the field of tennis modelling.

Comments: Comment Comment Commennt