Machine Learning and Its Application in Fantasy Football

Jonathan R. DiQuattro

Thomas Edison State University

Abstract

This paper is a draft chapter 2 of a research paper into the application of using machine learning to determine the top running backs for each season of the last decade with a 60% accuracy. This chapter primarily focuses on the references required for that research.

Keywords: Machine learning, fantasy football, algorithm.

Machine learning is the pinnacle of what computer science can achieve (Sharma, B.), and allows for the processing of an incredible almost unfathomable amounts of data into useful information, and has the potential of completely altering mankind. With such potential comes great complexity, and with great complexity derives a need for explaining concepts in a high level of of abstraction that can be filtered into more and more complex ideas. This chapter gathers the materials and sources needed to understand some of the the underlying theories of machine learning to better understand the problem domains, and make machine learning a topic that anyone outside of computer science can understand.

Body paragraph

With a topic like machine learning that requires a high level abstract view for both the seasoned computer scientist and the general population, it is important to understand the underlying proof, so that the abstract conceptualization is rooted in actual science. Some of the concepts in this paper are based on concrete math such as in a discrete mathematics (Kenneth H.) course or other high level maths to include linear algebra and vector spaces (Beezer, A. Robert). This paper will not be conducting any mathematical proofs, but the concepts and theories presented in this paper can be proven using advanced graphing techniques and Vector Spaces.

With the vast amounts of exponential data being created each year (Mearian, L.), machine learning techniques have been developed to handle the complex relationships of some times vast qualitative data inputs. These sometimes intricately connected data points create a number of problem domains (Segaran, Toby), this paper will primarily focus on the solution for 4 to six variable (or nodes) with multiple degrees of distribution(Kenneth H.), comprised of

team, coaching traits of hundreds of NFL coaches, and positional data for the last 20 years for 32 National Football League (NFL) teams.

One of the main questions this paper aims to answer is the possibility of calculating the data for the purpose of selecting a top three running back in each season of the last decade using the previous 20 years of NFL data, the answer to that question depended heavily on understanding Big O notation and algorithm complexity (Dale, Nell).

While the underlying concepts of machine learning are very complex, the tools used in machine learning have increased over the last decade, and today there are many frameworks available (Shaleynikov, Anton.), this research relied on tensorflow and the python coding language. While practically anyone who knows how to code can utilize machine learning, the difficulty in achieving an accuracy of over 60%, relied heavily on selecting the input variables that would carry the most weight (Segaran, Toby.).

One of the most difficult questions to answer was what machine learning model would be the best option for computing multiple degrees (potentially many degrees) of distribution, after considering a neural network (Gupta, D.)(Segaran, Toby.), it was determined that a k-NN (k-Nearest Neighbor), would be able to compute the large number degree of distribution in an acceptable amount of time, this was determined using Big O notation (Kenneth H.), to keep this paper at a high level of abstraction, Big O will not be explained, but was nonetheless calculated to be O(n)^d where n is the amount of operations and d is degrees of distribution.

The accuracy of k-NN in conjunction with many variables is heavily reliant upon the chosen inputs (Srivastava, Tavish.), coincidentally for this study the inputs align with football strategy, for example the outcome of the algorithm can be changed by increasing the weight that

rushing attempts has in the algorithm, which rushing attempts is typically a strong indication of how a team performs overall, one of the goals of this study is to shed some light on traditional strategy (Berry, Mathew.), and possibly provide future insight to anyone who plays fantasy football.

Conclusion

There is a lot of complexity when dealing with machine learning, that legitimately takes years and even lifetimes to learn, but thankfully the actual application of machine learning can be learned in a few weeks through high level abstraction (Smola, Alex. and Vishwanathan, S.V.N.), however access to the overarching publications is crucial for the success of any machine learning algorithm.

Annotated Bibliography

Mearian, L. By 2020, there will be 5,200 GB of data for every person on Earth. https://www.computerworld.com/article/2493701/data-center/by-2020--there-will-be-5-200-gb-of-data-for-every-person-on-earth.html. December 11, 2012. Web. Retrieved July 11.

This Article explains the exponential growth of data each successive year, and how machine learning allows for turning that data into usable information.

Sharma, B. 'Machine learning is the big one': Deloitte expert looks to the future of tech trends. https://www.cnbc.com/2018/04/10/machine-learning-is-the-big-one-deloitte-expert.html. April 10, 2018. Web. Retrieved July 4, 2018.

In this interview a top expert at Deloitte (a large research and development firm) talks about how many companies are investing in machine learning for enterprise reasons.

Gupta, D. Fundamentals of Deep Learning – Introduction to Recurrent Neural Networks. https://www.analyticsvidhya.com/blog/2017/12/introduction-to-recurrent-neuraL-networks/. December 7, 2017. Web. Retrieved July 8, 2018.

This blog explains how a recurrent neural network works, and some of its applications.

Beezer, A. Robert. A First Course in Linear Algebra. Gig Harbor, Washington. Congruent Press.

This textbook explains some of the complex math that is the basis for machine learning algorithms.

Segaran, Toby. (2008) Programming Collective Intelligence. Sebastopol, California, O'Reilly Publishing.

Collective Intelligence is considered by many people one of the most essential publications to understand the basics of machine learning, and was essential for the success of this study.

Dale, Nell. (2013) C++ Plus Data Structures. Burlington, MA, Jones & Bartlett Learning.

This is a C++ programming architecture book, and explains some common data structures and big O notation, and was useful in determining the efficiency of a K-NN algorithm.

Kenneth H. Rosen. Discrete Mathematics and Its Application. New York, NY, McGraw Hill.

This text book represents what one might learn in a discrete mathematic class, and is heavily applied in machine learning for calculating probability, big O, degree of distribution, and other complex graphing theory.

Smola, Alex. and Vishwanathan, S.V.N. Introduction to Machine Learning. New York, NY, Cambridge University Press.

Shaleynikov, Anton. 10 Best Frameworks and LibrariesforAI.https://dzone.com/articles /progressive-tools10-best-frameworks-and-libraries. January 10, 2018.Web. Retrieved June 28, 2018.

This article points out the many frameworks that can be used for machine learning, and provides a good place to start when considering how to actually implement a machine learning algorithm.

Srivastava, Tavish. Introduction to k-Nearest Neighbors: Simplified (with implementation in Python). https://www.analyticsvidhya.com/blog/2018/03/introduction-k-Neighbours-algorithm-clustering/. March 26 2018. Web. Retrieved July 28, 2018.

This blog introduces k-Nearest Neighbor and how you can use it for clustering which is a useful technique for data scientists.

Berry, Matthew. The Draft-Day Manifesto: Why a weekly focus is key to fantasy footballsuccess.httpt//www.espn.com/fantasy/football/story/_/page/TMRDraftDayManifesto18/strategy-risk-Management-historical-trends-tips-how-draft-fantasy-football. July 26, 2018. Web. Retrieved July 10, 2018.

Trusted by millions of fantasy football fans for two decades, Matthew Berry is a leading expert on trends and strategies used to win fantasy football leagues around the country each year.