NFL Score Difference Prediction with Markov Modeling

Guy Blanc, Eric Luxenberg, Stanley Xie
gblanc@stanford.edu, ericlux@stanford.edu, stanleyx@stanford.edu

Introduction

With discrete plays and clear features like score, down, time left, and a natural goal of predicting the winner and score of a game, American football possesses an unique and attractive structure for Machine Learning analysis. Our objective in this project is to predict the final score difference of a football game given three quarters of in-game information. We used an aggregated dataset with play-by-play data for the 2002-2013 NFL seasons, containing 467,199 plays over 2909 games.

Models/Algorithms

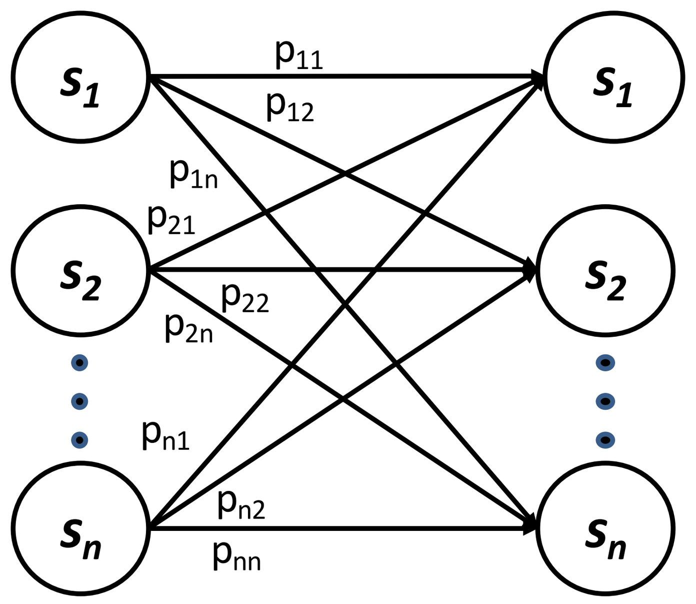


Figure 1. One time step of our Markov Model. States are determined by down, yardline, yard to go, average yards, plays left. Special states are assigned to scoring plays. The model runs based on game specific prediction of number of 4th quarter plays (see Figure 2.).

$$y^i \sim N(\theta_1 x^i, (\theta_2 x^i)^2)$$

$$\frac{\partial LL(\theta_1,\theta_2)}{\partial \theta_1}_j = \sum_{i=1}^{m} \frac{(y^i - \theta_1 x^i)x^i_j}{(\theta_2 x^i)^2}$$

$$\frac{\partial LL(\theta_1,\theta_2)}{\partial \theta_2}_j = \sum_{i=1}^{m} \frac{x^i_j((y^i - \theta_1 x^i)^2 - (\theta_2 x^i)^2}{(\theta_2 x^i)^3}$$

Figure 2. Estimating number of plays to model: M.L.E gradient descent rules assuming the number 4th quarter plays is normally distributed. with μ and σ linear in features from first three quarters. Game dependent estimation of the number of 4th quarter plays allows for more context specificity in our model.

Models/Algorithms cont.

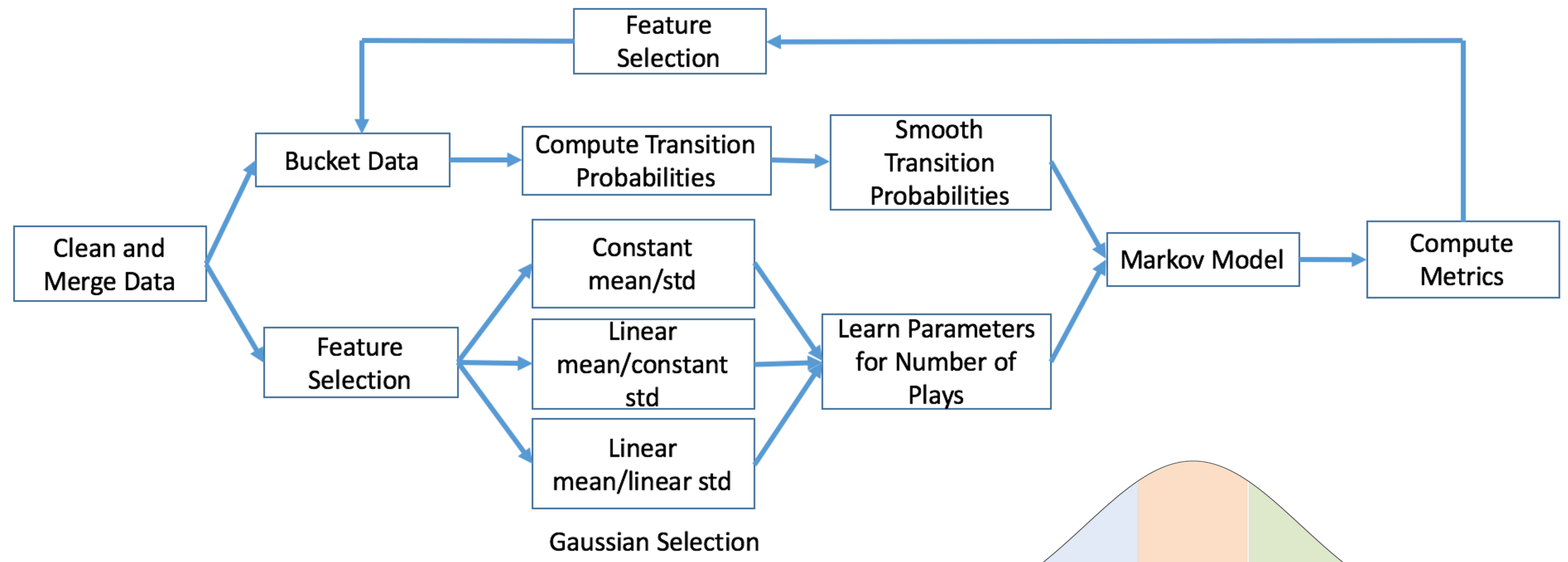


Figure 3. Our pipeline from data acquisition to prediction/scoring

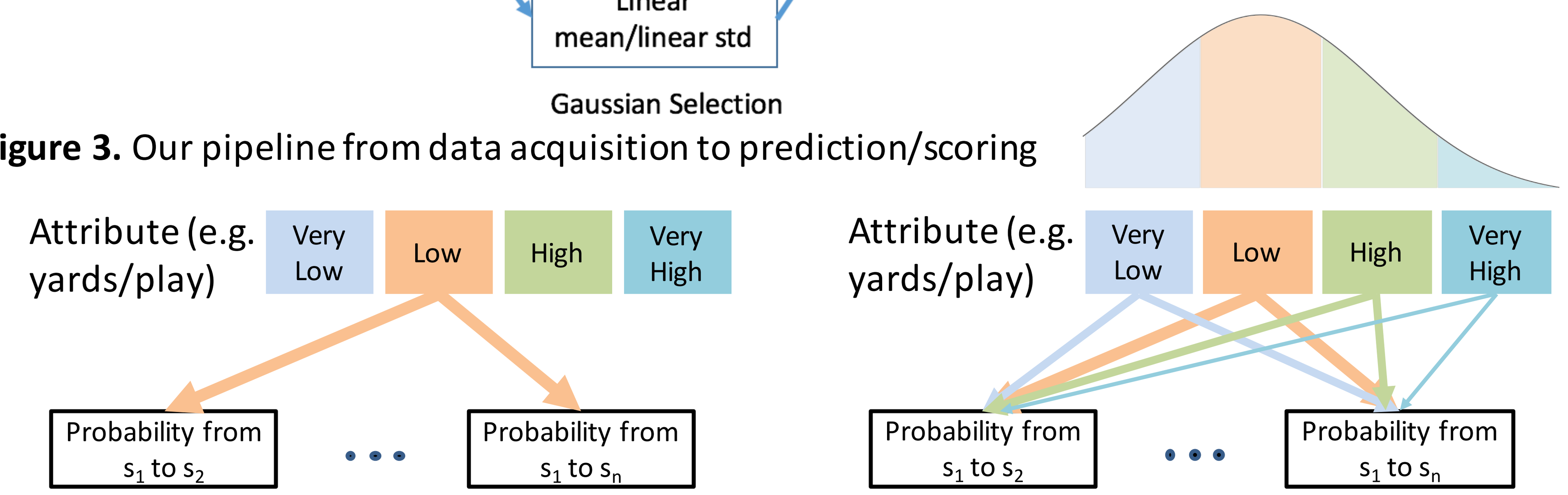
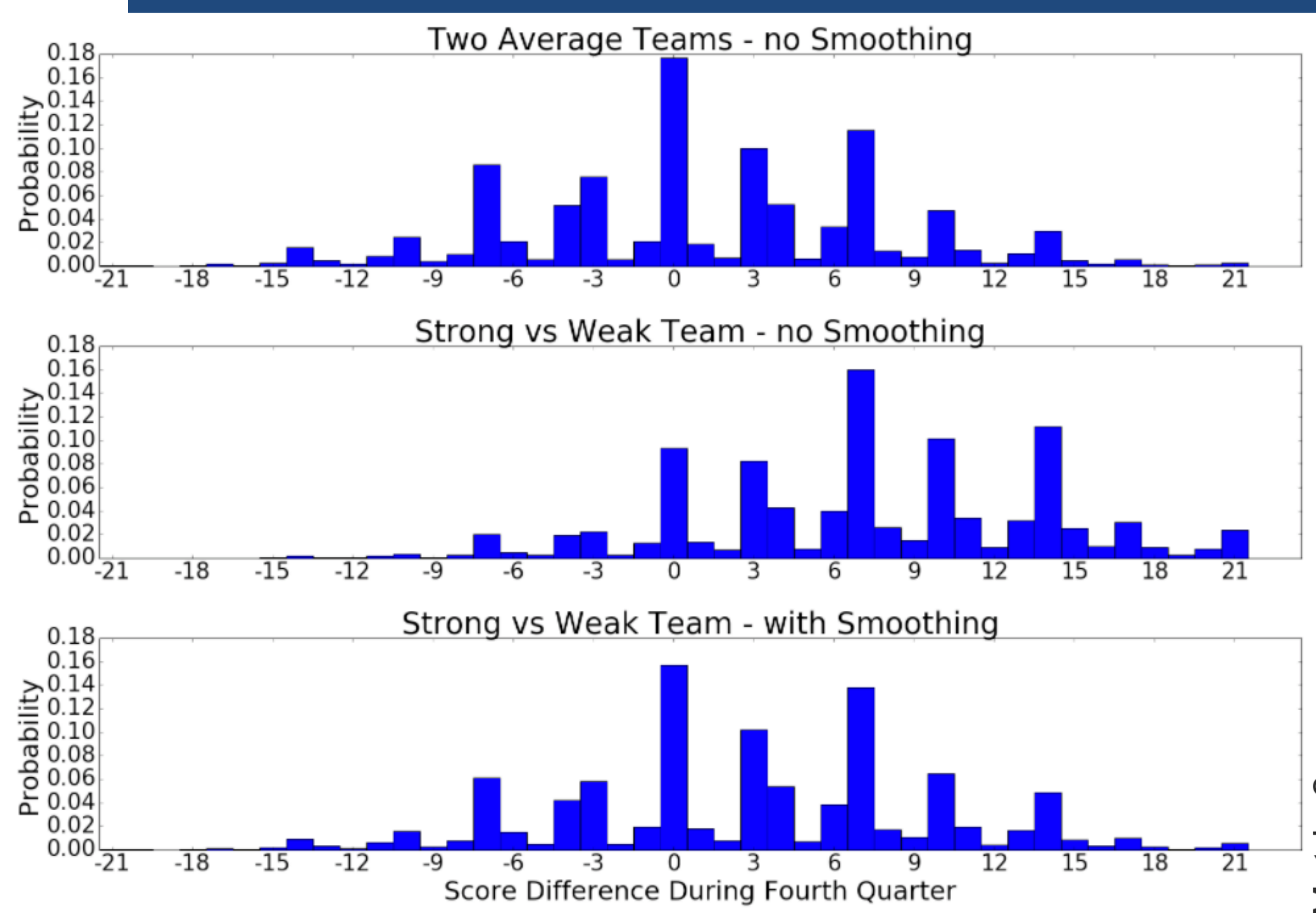


Figure 4. An illustration of smoothing along one parameter (average yards/play) using a Gaussian distribution

Results/Discussion



Metric 1: Average squared error between expected value of score diff. and actual score diff.
Metric 2: The mean log likelihood assigned to the actual score difference over all games
Metric 3: % of games where actual score diff. was in our top 5 score diffs (ranked by probability)

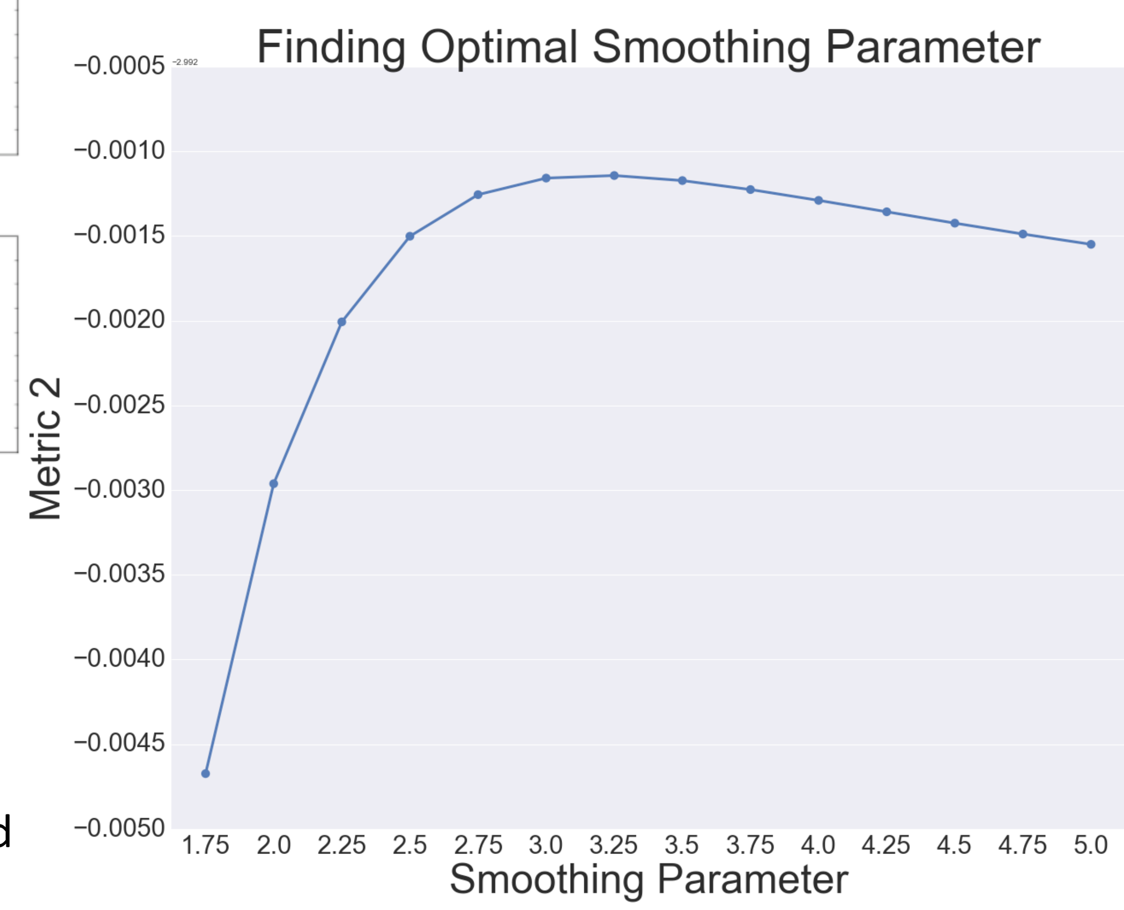


Figure 5. Probability densities for 4th quarter score difference with increasing model complexities. Smoothing causing reversion to mean.

Figure 6. Tuning the smoothing parameters for average yards per a play. At the limit of very large tuning parameter (standard deviation), average yards per play does not contribute to transition probabilities

Results/Discussion cont.

	Metric 1	Metric 2	Metric 3
Naïve	66.1013	-3.5145	0.2687
Position state, all quarters	51.2583	-3.0165	0.5374
Position state, 4th quarter	**51.0961**	-2.9934	**0.5408**
Extra features no smoothing	56.1134	-3.0152	0.5136
Extra features/smoothing	51.2906	-2.9698	0.5374
Extra features/smoothing with num. plays est.	51.3062	**-2.9692**	0.5374

Table 1. Performance of various iterations of our model as measured by our three metrics.

- We found that training on only 4^{th} quarter data improved our results, likely because teams play differently in the 4th quarter than the rest of the game (e.g. if a team is winning by a large margin, it likely will slow down its scoring).
- The most predictive features for the number of plays in the fourth quarter outperformed the test and train likelihood of a naive normal distribution with sample mean and sample variance.
- Smoothing and MLE estimations of number of plays in the fourth quarter improved our results

Conclusion/Future Work

Our model provided a predictive edge for end of game score difference. Insights generated included fourth quarter plays estimation and transition probability smoothing. In the future, we would like to utilize even more in-game information, perhaps incorporating previous plays into our models. Furthermore, we can start this model at different points in the game and use it to provide an estimate of in-game win probability.

References

1. Gimpel, K. (2006) *Beating the NFL football point spread*. Available at: http://ttic.uchicago.edu/~kgimpel/papers/machine-learning-project-2006.pdf (Accessed: 12 November 2016).
2. Glickman, M.E. and Stern, H.S. (1998) 'A state-space model for national football league scores', *Journal of the American Statistical Association*, 93(441), p. 25. doi: 10.2307/2669599.
3. Glickman, M.E. and Stern, H.S. (2016) *Estimating team strength in the NFL*. Available at: http://www.glicko.net/research/nfl-chapter.pdf (Accessed: 19 November 2016).
4. Hirotsu, N. and Wright, M. (2002) 'Using a Markov process model of an association football match to determine the optimal timing of substitution and tactical decisions', *The Journal of the Operational Research Society*, 53(1), pp. 88–96. doi: 10.2307/822882.
5. Matthews, G.J., Wilbur, J.D. and Vernescu, B.M. (2005) *Improving paired comparison models for NFL point spreads by data transformation*. Available at: https://web.wpi.edu/Pubs/ETD/Available/etd-050505-144934/unrestricted/Thesis.pdf (Accessed: 12 November 2016).
6. Peña, J.L. (2014) *A Markovian model for association football possession and its outcomes*. Available at: https://arxiv.org/pdf/1403.7993.pdf (Accessed: 1 December 2016).
7. Warner, J. (2010) *Predicting Margin of Victory in NFL Games: Machine Learning vs. The Las Vegas Line*. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.402.8415&rep=rep1&type=pdf (Accessed: 28 October 2016).