Rishi Mody

Machine Learning
CS 589

MS UMass CS, 2016-18

Graduate Research Intern, UMass Amherst
Software Engineering Intern, Nectar Globe Technology Solutions, India
Btech ‘16, ECE, VNIT Nagpur, India

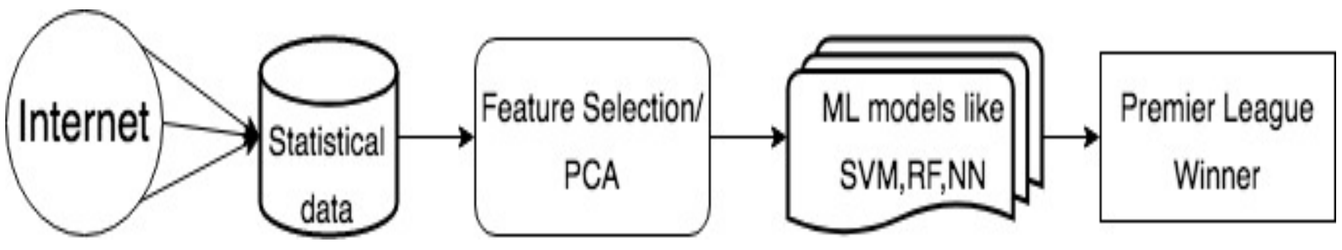
Machine Learning to Predict The English Premier League Winner

Problem

- Forecasting league winners on the basis of previous years’ data.
- Identifying aspects of a team’s gameplay which affect game results the most.

Solution

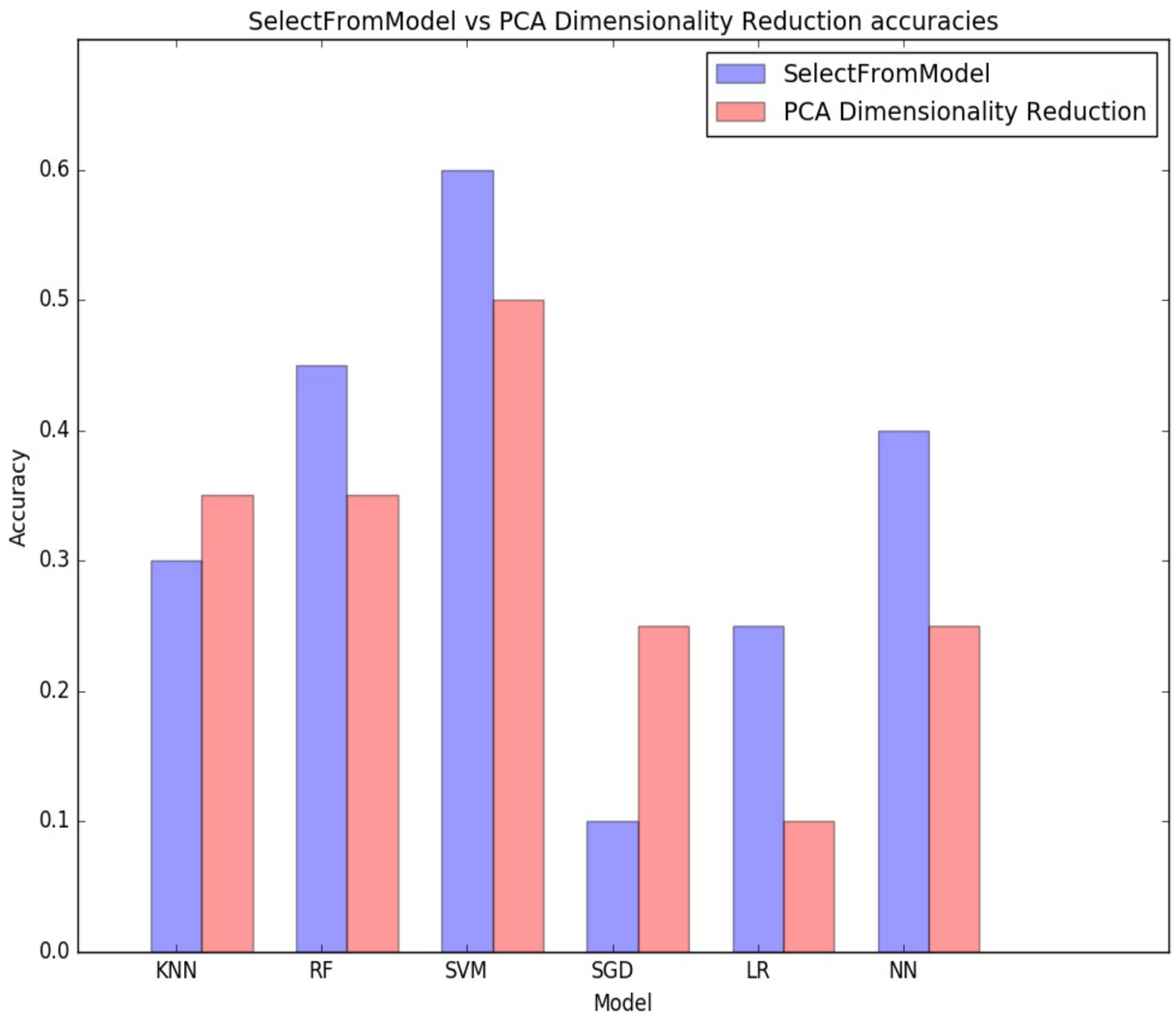
- Feature Engineering - Collate data from internet sources of not only game results per team for multiple years, but also match stats including goals scored, fouls committed, bookings etc.
- Feature Selection/PCA to focus on features that impact the result the most.
- Evaluation and analysis of various ML models for prediction across play seasons.



$$f_{SVM}(x) = sign(w^T x + b) \qquad \hat{f} = \frac{1}{B}\sum_{b=1}^{B} \hat{f}_b(\acute{x})$$

Results

		Accuracies after Feature Selection					
Trained on seasons(starting years)	Tested on seasons	KNN	RF	SVM	SGD	LR	Neural Network
2003	2004	0.25	0.25	0.1	0.25	0.15	0.3
2003-2004	2005	0.25	0.3	0.3	0.1	0.2	0.3
2003-2005	2006	0.45	0.45	0.4	0.2	0.1	0.35
2003-2006	2007	0.4	0.45	0.35	0.1	0.2	0.25
2003-2007	2008	0.4	0.55	0.5	0.35	0.35	0.5
2003-2008	2009	0.35	0.4	0.2	0.1	0.1	0.2
2003-2009	2010	0.3	0.4	0.35	0.15	0.25	0.35
2003-2010	2011	0.3	0.45	0.6	0.1	0.25	0.45
2003-2011	2012	0.35	0.5	0.55	0.35	0.2	0.4
2003-2012	2013	0.4	0.5	0.5	0.1	0.15	0.5
2003-2013	2014	0.2	0.5	0.35	0.1	0.1	0.5



Future work

Predict the expected number of goals scored in a match depending on the previous matches of the season.

Coursework

ML, Neural Networks, NLP, Information Retrieval, Advanced Algos, Advanced Software Engineering, Databases