Bayesian Networks
AI Assignment - 3

CSK vs RCB

Team 63

Abhinav Gupta - 20171059

Praffullitt Jain - 20171142

Modelling a Bayes' Net for a Cricket Match

The Question

"It being IPL season, you have registered yourself for a fantasy league which earns you points if you predict the result of the match correctly. It is the match between CSK and RCB at the Chinnaswamy stadium(Bangalore). Model a bayesian network predicting who would win the game considering factors like team strength, weather conditions, home ground etc. Give conditional probabilities and justify."

The Variables

So first off, let us think of all the nodes that would be present in our Bayes network.

Layer 1:

- The simplest - 2 nodes "RCB" and "CSK", which would indicate which team wins the match. Now we need to think, on what factors does a CSK win or an RCB win depend on?

Layer 2:

- It would definitely be a win for that team which has really strong players! So the quality, ICC rating and experience of players would definitely play an important role in deciding who wins the match. For instance, a team of 3 consisting of MS Dhoni, Steve Smith and Dale Steyn would definitely be stronger than that of say, Tim Paine, Stuart Broad and Ajinkya Rahane….because the former team has better statistics, better averages and better rating than the latter team. So we can have a variable as TEAM STRENGTH which would determine quality of players.
- Another factor affecting a win or loss would be - CONFIDENCE! It's extremely important for the players to be confident and play fearlessly.

- And finally - the PITCH! Before very cricket match, a lot of time is spent in studying the pitch and doing a thorough analysis of it. Why? Because pitch conditions really matter. For instance, a moist pitch would really be useful to spin bowlers as it would be easy to baffle the batsman.

Layer 3:

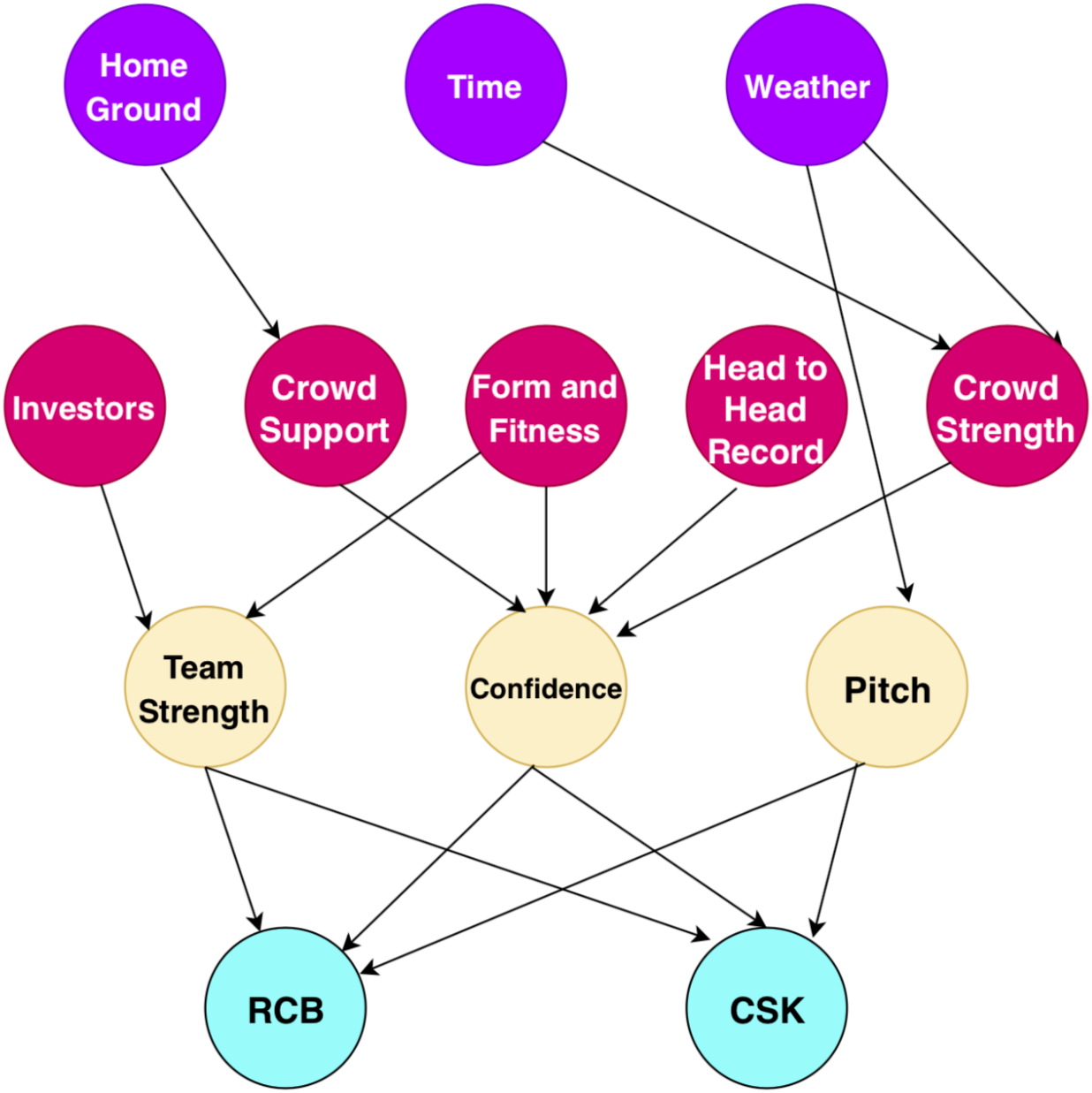
- The confidence of the team would further depend on factors like CROWD! For instance, if the match is being played on home ground of CSK, then CSK will have overwhelming support from the crowd. That way, everyone would be cheering for players in the yellow jersey, which would be a great boost to their confidence. On the other hand, everyone would be booing for RCB, which will lower their confidence. Maybe some people may not agree with 'confidence', but it really is a important factor. After all, even in football, away goals are considered more valuable than home goals because it is harder to score when crowd is not on your side!
- Confidence will also depend on the FORM of the players! Form means how well the player has been doing recently…like for instance, the performance of the player in the 10 games prior to the current game. If the player has had a very good average in the games just prior to the current one, he/she is definitely in good form and this would be a major confidence booster.
- Confidence would be highly dependent on the head to head record of the teams. If a team has beaten the other many times before, it's a pretty good confidence booster.
- Team Strength would also depend on form and fitness. If there are more players who are in form and playing really well, it would certainly increase the team's strength.
- Team Strength would depend on INVESTORS, whether they have enough money to be able to afford to buy quality players for the team. For instance, MS Dhoni would definitely cost more than Rs. 10 crore, so the investors must be able to afford that much.

Layer 4:

- The Crowd Support would depend on the HOME GROUND of the match, whether it is in home ground of CSK or RCB. The crowd strength would depend on WEATHER conditions, like people would not prefer to come for match if it is unbearably hot. Similarly, it would depend on TIME. If the

match is being held early in the morning at 7am, there will be lesser crowd people would not prefer to wake up so early.

- Pitch will also be dependent on the weather, whether it's sunny or rainy.



Key:

RCB	Victory for RCB	Yes, No
RCB	Victory for RCB	Win, Lose
CSK	Victory for CSK	Win, Lose
TS	Team Strength	Strong, Average, Weak
C	Confidence	GoodforCSK, GoodforRCB
H2H	Head to Head Record	CSK, RCB (have won more matches)
P	Pitch	Good for Batting, Good for Bowling
FF	Form and Fitness	Good, Bad
CWP	Crowd Support	GoodforCSK, GoodforRCB
CWS	Crowd Strength	High, Low
I	Investors	Rich, Average
HG	Home Ground	RCB, CSK
W	Weather	Rainy, Hot, Other
T	Time	Day, Night

Conditional Probability Tables:

Home Ground:

HG	RCB	CSK
P(HG)	0.5	0.5

Of course, because in cricket there are two legs off matches, one at each of their homegrounds. So in this case, one match will be played in RCB's playground and one in CSK's playground.

Weather:

W	Rainy	Hot	Other
P(W)	0.2	0.5	0.3

Approximating these values after studying the temperatures of Bangalore and Chennai. In both places, most of the time it is summer, and rains pretty often. 'Other' would be the ideal temperature. No one wants to sit in the hot sunny weather in the open stadiums.

Time:

T	Day	Night
P(T)	0.2	0.8

Given latest IPL statistics, almost all the matches are in the night time, because better crowd. But sometimes, due to unforeseen circumstances like bad weather, and if the match is really important, the match may get postponed to the next morning; hence the probability of 0.2.

Crowd Support:

Home Ground (HG)	CWP.GoodforCSK	CWP.GoodforRCB
RCB	0.2	0.8
CSK	0.8	0.2

If the match is in home ground of CSK, then there will be more supporters of CSK than RCB. Hence, lots of support for them, and vice-versa.

Crowd Strength:

This is different from crowd support. It also makes a difference as to how many people have come to watch the match. If the crowd is high and lots of people are there in the stadium, there is a greater feeling of excitement and enthusiasm in the air, boosting confidence of the players.

Time(T)	Weather(W)	CWS.High	CWS.Low
Day	Other	0.7	0.3
Day	Hot	0.5	0.5
Day	Rainy	0.2	0.8
Night	Other	0.9	0.1
Night	Hot	0.7	0.3
Night	Rainy	0.4	0.6

Naturally people would prefer to go for the match in the night time, as it is more convenient compared to day time, where people may have to work. And people would most certainly avoid rainy weather and hot weather, because it becomes hard to sit in such weather in the open stadium. Normal/other weather would be preferred.

Form and Fitness

FF	Good	Bad
P(FF)	0.8	0.2

Usually players are in form, and fit to play. But injuries or bad phase can come for even the best of cricketers, resulting in the 0.2 probability for bad form.

Investors

I	Rich	Average
P(I)	0.3	0.7

Rich investors would mean the Ambani's or the Tata's. They would have millions of dollars to spend to buy quality players and spend for coaching and training. But there are less 'very rich' investors in IPL, so the probability is lower.

Pitch

Weather	P.GoodforBatting	P.GoodforBowling
Hot	0.8	0.2
Rainy	0.3	0.7
Other	0.5	0.5

Pitch conditions are usually kept normal and standard, with some amount of grass cover and appropriate amounts of dampness or dryness.
But when it rains, pitch becomes damper with some grass cover, making it better for bowling. Dampness helps the spinners, bamboozling the batsman.
However, dryer pitches with lesser grass cover in the hot weather would not affect the trajectory of the ball after bouncing. So batsmen can easily predict the length of the delivery, and play accordingly. Hence it becomes better for batting.
Source: https://www.quora.com/What-is-the-difference-between-batting-and-bowling-pitch-in-cricket

Head to Head Record

H2H	RCB	CSK
P(H2H)	0.3	0.7

Because as per prior statistics, CSK has won more matches and has a better record against RCB.

Team Strength

Investors (I)	Form and Fitness (FF)	TS.Strong	TS.Average	TS.Weak
Rich	Good	0.7	0.2	0.1
Rich	Bad	0.4	0.35	0.25
Average	Good	0.5	0.3	0.2
Average	Bad	0.3	0.4	0.3

If the investors are rich, they are able to afford to buy expensive quality players for more money, thereby increasing the team strength. And of course, if the players are not injured and in form, then definitely, it will increase team strength!

Confidence

Crowd Support (CWP)	Crowd Strength (CWS)	Form and Fitness (FF)	Head to Head Record (H2H)	C.Goodfor CSK	C.Goodfor RCB
GoodforCSK	High	Good	CSK	0.95	0.05
GoodforCSK	High	Good	RCB	0.9	0.1
GoodforCSK	High	Bad	CSK	0.8	0.2
GoodforCSK	High	Bad	RCB	0.7	0.3
GoodforCSK	Low	Good	CSK	0.9	0.1
GoodforCSK	Low	Good	RCB	0.8	0.2
GoodforCSK	Low	Bad	CSK	0.85	0.15
GoodforCSK	Low	Bad	RCB	0.65	0.35
GoodforRCB	High	Good	CSK	0.1	0.9
GoodforRCB	High	Good	RCB	0.05	0.95
GoodforRCB	High	Bad	CSK	0.3	0.7
GoodforRCB	High	Bad	RCB	0.2	0.8

GoodforRCB	Low	Good	CSK	0.2	0.8
GoodforRCB	Low	Good	RCB	0.1	0.9
GoodforRCB	Low	Bad	CSK	0.35	0.65
GoodforRCB	Low	Bad	RCB	0.15	0.85

There are a lot of factors on which confidence depends!

- Crowd support is of paramount importance; if the match is in RCB's home-ground, then definitely RCB will have great support from the crowd which will boost its confidence tremendously.
- If crowd strength in general is high, there will be more confidence in all players as lots of people have come to watch them.
- Head to head record matters as well. If CSK knows that they've beaten RCB more times before, it certainly makes them feel better.
- And finally, if form and fitness of a team is good, they will naturally feel better about themselves, and have faith. If the players are in good form, they will be more confident.

RCB

We assume that CSK has a better batting team and RCB has a better bowling team. So Pitch will accordingly influence their winning chances.

Team Strength(TS)	Confidence(C)	Pitch(P)	RCB.Win	RCB.Lose
Strong	GoodforCSK	Goodfor Batting	0.5	0.5
Strong	GoodforCSK	Goodfor Bowling	0.7	0.3
Strong	GoodforRCB	Goodfor Batting	0.7	0.3
Strong	GoodforRCB	Goodfor Bowling	0.9	0.1
Average	GoodforCSK	Goodfor Batting	0.35	0.65

Average	GoodforCSK	Goodfor Bowling	0.55	0.45
Average	GoodforRCB	Goodfor Batting	0.55	0.45
Average	GoodforRCB	Goodfor Bowling	0.75	0.25
Weak	GoodforCSK	Goodfor Batting	0.2	0.8
Weak	GoodforCSK	Goodfor Bowling	0.4	0.6
Weak	GoodforRCB	Goodfor Batting	0.4	0.6
Weak	GoodforRCB	Goodfor Bowling	0.55	0.45

CSK

We assume that CSK has a better batting team and RCB has a better bowling team. So Pitch will accordingly influence their winning chances.

Team Strength(TS)	Confidence(C)	Pitch(P)	CSK.Win	CSK.Lose
Strong	GoodforCSK	Goodfor Batting	0.9	0.1
Strong	GoodforCSK	Goodfor Bowling	0.7	0.3
Strong	GoodforRCB	Goodfor Batting	0.7	0.3
Strong	GoodforRCB	Goodfor Bowling	0.5	0.5
Average	GoodforCSK	Goodfor Batting	0.75	0.25

Average	GoodforCSK	Goodfor Bowling	0.55	0.45
Average	GoodforRCB	Goodfor Batting	0.55	0.45
Average	GoodforRCB	Goodfor Bowling	0.35	0.65
Weak	GoodforCSK	Goodfor Batting	0.55	0.45
Weak	GoodforCSK	Goodfor Bowling	0.4	0.6
Weak	GoodforRCB	Goodfor Batting	0.4	0.6
Weak	GoodforRCB	Goodfor Bowling	0.2	0.8

Sample Query:

Let us take the query of the form P(X | p(X), p(p(X)))
Where x is the variable
p(X) is its parent
p(p(X)) is the parent of its parent.

We want to find:

$$P(RCB = win \mid Pitch=GoodforBowling, Weather = rainy)$$

Solution:

So we visit table of RCB. We fix Pitch = GoodforBowling, and its probability becomes 1. That is, we neglect all other rows where Pitch is GoodforBatting.

RCB
RCB.Win * Confidence * Team Strength * (Pitch = GoodforBowling)

So now we need to find

P(Confidence = GoodforCSK)
P(Confidence = GoodforRCB)
P(TS = Strong)
P(TS = Average)
P(TS = Weak)

Let's find these one by one.

P(Confidence = GoodforCSK)
C.GoodforCSK * CWS * CWP * FF * H2H

So this way, we keep recursing upto the parents, until we find our final solution.

So now we need to find:

P(H2H = CSK) = 0.7
P(H2H = RCB) = 0.3
P(FF = Good) = 0.8
P(FF = Bad) = 0.2
P(CWS = Low)
P(CWS = High)
P(CWP = GoodforCSK)
P(CWP = GoodforRCB)

P(CWP = GoodforCSK) = CWP.GoodforCSK * HG
So we need to find P(HG)

P(HG = RCB) = 0.5
P(HG = CSK) = 0.5

P(CWP = GoodforCSK) = CWP.GoodforCSK * HG
(0.2*0.5) + (0.8*0.5) = 0.5

P(CWP = GoodforRCB) = CWP.GoodforRCB * HG
(0.8 * 0.5) + (0.2*0.5) = 0.5

P(CWS = High) = CWS.High * Time * Weather=Rainy
= (0.2 * time=day) + (0.4 * time = night)

P(time = day) = 0.2
P(time = night) = 0.8

P(CWS=High) = 0.2 * 0.2 + 0.4 * 0.8 = 0.04 + 0.32 = 0.36

P(CWS = Low) = CWS.Low * Time * Weather=Rainy
0.8 * 0.2 + 0.6 * 0.8 = 0.16 + 0.48 = 0.64

So the values are:
P(H2H = CSK) = 0.7
P(H2H = RCB) = 0.3
P(FF = Good) = 0.8
P(FF = Bad) = 0.2
P(CWS = Low) = 0.64
P(CWS = High) = 0.36
P(CWP = GoodforCSK) = 0.5
P(CWP = GoodforRCB) = 0.5

P(Confidence = GoodforCSK)
C.GoodforCSK * CWS * CWP * FF * H2H

0.5	0.36	0.8	0.7	0.95 = 0.09576
0.5	0.36	0.8	0.3	0.9 = 0.03888
0.5	0.36	0.2	0.7	0.8 = 0.02016
0.5	0.36	0.2	0.3	0.7 = 0.00756
0.5	0.64	0.8	0.7	0.9 = 0.16128
0.5	0.64	0.8	0.3	0.8 = 0.06144
0.5	0.64	0.2	0.7	0.85 = 0.03808
0.5	0.64	0.2	0.3	0.65 = 0.01248
0.5	0.36	0.8	0.7	0.1 = 0.01008
0.5	0.36	0.8	0.3	0.05 = 0.00216
0.5	0.36	0.2	0.7	0.3 = 0.00756
0.5	0.36	0.2	0.3	0.2 = 0.00216
0.5	0.64	0.8	0.7	0.2 = 0.03584
0.5	0.64	0.8	0.3	0.1 = 0.00768

0.5 0.64 0.2 0.7 0.35 = 0.01568
0.5 0.64 0.2 0.3 0.15 = 0.00288

P(Confidence = GoodforCSK) = 0.51968

P(Confidence = GoodforRCB)
C.GoodforRCB * CWS * CWP * FF * H2H

Table 1

	0.05000	0.7	0.8	0.36	0.5	0.00504
	0.1	0.30000	0.8	0.36	0.5	0.00432
	0.2	0.70000	0.2	0.36	0.5	0.00504
	0.3	0.30000	0.2	0.36	0.5	0.00324
	0.1	0.70000	0.8	0.64	0.5	0.01792
	0.2	0.30000	0.8	0.64	0.5	0.01536
	0.15	0.70000	0.2	0.64	0.5	0.00672
	0.35	0.30000	0.2	0.64	0.5	0.00672
	0.9	0.70000	0.8	0.36	0.5	0.09072
	0.95	0.30000	0.8	0.36	0.5	0.04104
	0.7	0.70000	0.2	0.36	0.5	0.01764
	0.8	0.30000	0.2	0.36	0.5	0.00864
	0.8	0.70000	0.8	0.64	0.5	0.14336
	0.9	0.30000	0.8	0.64	0.5	0.06912
	0.65	0.70000	0.2	0.64	0.5	0.02912
	0.85	0.30000	0.2	0.64	0.5	0.01632
						0.48032

P(Confidence = GoodforRCB) = 0.48032

P(TS = Strong) = 0.514

0.7	0.80000	0.3	0.168
0.4	0.2	0.30000	0.024
0.5	0.8	0.70000	0.28
0.3	0.2	0.70000	0.042
			0.514

P(TS = Average) = 0.293

0.2	0.80000	0.3	0.048
0.35	0.2	0.30000	0.021
0.3	0.8	0.70000	0.168
0.4	0.2	0.70000	0.056
			0.293

P(TS = Weak) = 0.193

0.1	0.80000	0.3	0.024
0.25	0.2	0.30000	0.015
0.2	0.8	0.70000	0.112
0.3	0.2	0.70000	0.042
			0.193

RCB.Win * Confidence * Team Strength * (Pitch = GoodforBowling)

P(RCB = Win | Pitch = GoodforBowling, Weather = rainy) = 0.689578912

0.514	0.51968	0.7	0.186980864	
0.514	0.48032	0.9	0.222196032	
0.293	0.51968	0.55	0.083746432	
0.293	0.48032	0.75	0.10555032	
0.193	0.51968	0.4	0.040119296	
0.193	0.48032	0.55	0.050985968	
			0.689578912	

Hence, P(RCB = win | Pitch=GoodforBowling, Weather = rainy) ~ **0.69**