

# **Interaction filtering - A novel approach to social recommendation**

**Riley Kidd**

A subthesis submitted in partial fulfillment of the degree of  
Bachelor of Software Engineering at  
The Department of Computer Science  
Australian National University

October 2012

© Riley Kidd

Typeset in Palatino by  $\text{\TeX}$  and  $\text{\LaTeX} 2_{\epsilon}$ .

Except where otherwise indicated, this thesis is my own original work.

Riley Kidd  
8 October 2012



---

# Abstract

---

Social networks provide a wide array of user specific interactions, profile information and user preferences. This thesis attempts to decipher which user traits are truly indicative of 'likes', this information is then leveraged to allow for binary classification of user specific links with the goal of discovering the ideal combination of traits for prediction.

The success of our predictions are evaluated using a number of machine learning algorithms including, *Naive Bayes*, *Logistic Regression* and *Support Vector Machines*, results are also compared to previous work using *Matchboxing* and *Social Matchboxing* techniques. The data set is sourced from a set of over 100 Facebook users and their interactions with over 30,000 friends during a four month period.

We have shown that...



---

# Contents

---

|   |           |
|---|-----------|
| <b>Abstract</b>                         | <b>v</b>  |
| <b>1 Introduction</b>                   | <b>1</b>  |
| 1.1 Objectives . . . . .                | 1         |
| 1.2 Contributions . . . . .             | 1         |
| 1.3 Outline . . . . .                   | 1         |
| <b>2 Background</b>                     | <b>3</b>  |
| 2.1 Social Networks . . . . .           | 3         |
| 2.2 Data Set . . . . .                  | 3         |
| 2.3 Notation . . . . .                  | 5         |
| 2.4 Feature Sets . . . . .              | 5         |
| 2.5 Previous Work . . . . .             | 5         |
| 2.5.1 Content Based Filtering . . . . . | 5         |
| 2.5.2 Information Diffusion . . . . .   | 5         |
| 2.6 Prediction Algorithms . . . . .     | 5         |
| 2.6.1 Constant . . . . .                | 5         |
| 2.6.2 Social Recommender . . . . .      | 6         |
| 2.6.3 Naive Bayes . . . . .             | 6         |
| 2.6.4 Logistic Regression . . . . .     | 6         |
| 2.6.5 Support Vector Machine . . . . .  | 6         |
| 2.7 Training and Testing . . . . .      | 6         |
| 2.8 Evaluation Metrics . . . . .        | 7         |
| <b>3 User Interactions</b>              | <b>9</b>  |
| 3.1 Introduction . . . . .              | 9         |
| 3.2 Interactions . . . . .              | 9         |
| 3.3 Conversation . . . . .              | 12        |
| 3.3.1 Outgoing . . . . .                | 14        |
| 3.3.2 Incoming . . . . .                | 14        |
| <b>4 User Preferences</b>               | <b>15</b> |
| 4.1 Introduction . . . . .              | 15        |
| 4.2 Demographics . . . . .              | 15        |
| 4.3 Traits . . . . .                    | 20        |
| 4.4 Groups . . . . .                    | 26        |
| 4.5 Pages . . . . .                     | 30        |

|          |                                 |           |
|----------|---------------------------------|-----------|
| <b>5</b> | <b>Bayesian Model Averaging</b> | <b>33</b> |
| 5.1      | Introduction . . . . .          | 33        |
| 5.2      | Derivation . . . . .            | 33        |
| 5.3      | Results . . . . .               | 33        |
| <b>6</b> | <b>Conclusions</b>              | <b>35</b> |
| 6.1      | Summary . . . . .               | 35        |
| 6.2      | Future Work . . . . .           | 35        |



# Introduction

---

## 1.1 Objectives

## 1.2 Contributions

## 1.3 Outline

The goal of this thesis is to discover which sub-set or combination of user interactions and/or user preferences will be the most predictive of user likes.



# Background

---

## 2.1 Social Networks

The social network central for this study is Facebook. Once registering, Facebook users have the option of setting up a personalised profile, they can then establish themselves as friends of other users. Friends can interact via wall posts, conversations or by liking some facebook element.

Social networks such as Facebook provide a wide array of user preferences (link, tag, photo, video likes) in an array of interaction mediums and modalities (outgoing, incoming) as well as user specific information (gender, age, location, group memberships, favorite movies) and conversation content.

A problem with the Facebook paradigm in relation to this analysis is the requirement for assumed dislikes, if a user does not like some link can we imply the user does not like this link? Given the time period Facebook shows a link and the differing online times for Facebook users, this is generally a poor assumption. As such a Facebook app named LinkR was developed by NICTA which explicitly stores like and dislike data for users. This app will be discussed in the following section.

## 2.2 Data Set

The LinkR Facebook app was used to collect information about users, their interactions and preferences. The data set contains information about app users as well as a sub-set of visible information about their friends. The app tracked and stored information for over 100 app users and their 39,000+ friends.

The four main interactions between users are posts (posting an element on a friends' wall), tags (being mentioned in a friends post or comment), comments (written data on a post) and likes (clicking a like button if a user likes a post or comment). The table below outlines data collected during app trials.

| <b>App Users</b>             | <b>Posts</b> | <b>Tags</b> | <b>Comments</b> | <b>Likes</b> |
|------------------------------|--------------|-------------|-----------------|--------------|
| <b>Wall</b>                  | 27,955       | 5,256       | 15,121          | 11,033       |
| <b>Link</b>                  | 3,974        | -           | 5,757           | 4,279        |
| <b>Photo</b>                 | 4,147        | 22,633      | 8,677           | 5,938        |
| <b>Video</b>                 | 211          | 2,105       | 1,687           | 710          |
| <b>App Users and Friends</b> | <b>Posts</b> | <b>Tags</b> | <b>Comments</b> | <b>Likes</b> |
| <b>Wall</b>                  | 3,384,740    | 912,687     | 2,152,321       | 1,555,225    |
| <b>Link</b>                  | 514,475      | -           | 693,930         | 666,631      |
| <b>Photo</b>                 | 1,098,679    | 8,407,822   | 2,978,635       | 1,960,138    |
| <b>Video</b>                 | 56,241       | 858,054     | 463,401         | 308,763      |

**Table 2.1:** Total app user records

## 2.3 Notation

The mathematical notation used by our Predictors during this thesis are outlined below.

- A dataset  $D$  comprised of  $N$  user feature vectors  $x$  of size  $I$ , where each element  $x_i \in \{0, 1\}$
- The feature size of  $I$  for each  $x$  is defined by the current feature set being tested.
- A mapping for each tested vector  $N$  from  $D$  of the form  $x \rightarrow y$  where  $y$  is the prediction for this feature vector of the form  $y \in \{0 \text{ (like)}, 1 \text{ (dislike)}\}$  based on each element in  $x_i$

## 2.4 Feature Sets

The feature sets in  $x$  can be any of the following, which are discussed further in :::

- Interactions
- Demographics
- Traits
- Groups
- Pages
- Outgoing Messages
- Incoming Messages

## 2.5 Previous Work

### 2.5.1 Content Based Filtering

### 2.5.2 Information Diffusion

## 2.6 Prediction Algorithms

This analysis makes use of the results from a number of different prediction algorithms which are outlined below.

### 2.6.1 Constant

The constant predictor returns a constant result irrespective of the feature vectors selected from above. The most common result in our data set is *False* and hence the *False* predictor is used in our analysis.

## 2.6.2 Social Recommender

## 2.6.3 Naive Bayes

*Naive Bayes* (NB) is a basic predictor which involves applying Bayes' theorem using independence assumptions between each feature in  $x$ .

The NB implementation used during this thesis is an implementation previously devised by *Scott Sanner* [?].

## 2.6.4 Logistic Regression

*Logistic Regression* (LR) predicts the odds of being either a like or a dislike by converting a dependent variable and one or more continuous independent variable(s) into probability odds.

The LR implementation used during this thesis is *LingPipe* [?], which supports:

- Three priors for regression (Cauchy, Gaussian, Laplace)
- Maximum entropy mode

## 2.6.5 Support Vector Machine

The *Support Vector Machine* (SVM) is a supervised learning machine based on a set of basis functions which help construct a separating hyperplane between the data points. Training involves building the relevant hyperplanes which can then be used for testing. Each data point is classified depending on which side of the hyperplane it falls.

The SVM implementation used during this thesis is *SVMLibLinear* [?], which supports:

- L2-regularized classifiers
- L2-loss linear SVM, L1-loss linear SVM, and logistic regression (LR)
- L1-regularized classifiers
- L2-loss linear SVM and logistic regression (LR)
- L2-regularized support vector regression
- L2-loss linear SVR and L1-loss linear SVR.
- Probability estimates

## 2.7 Training and Testing

All evaluation is done using 10 fold cross validation wherein the data is partitioned into 10 complimentary subsets, each subset is composed of two separate parts one section is used for training (80%) and the other (20%) is used for testing. This is performed on 10 distinct subsets and the results are averaged across each fold.

---

## 2.8 Evaluation Metrics

When evaluating the success of each method at correctly predicting the classification, the following metrics will be used.

- A *true positive* prediction refers to when the classifier correctly identifies the class as true.
- A *false positive* occurs when the prediction is true, but the true class was false.
- A *false negative* occurs when the prediction is false but the actual class is true.

Accuracy relates to the closeness of the true value. In the context of our results, the accuracy refers to the number of correct classifications divided by the size of the data set.

$$\text{accuracy} = \frac{\text{number of correct classifications}}{\text{size of the test data set}}$$

Precision relates to the number of retrieved predictions which are relevant. In the context of our results, the precision refers to the number of true positive predictions divided by the sum of the true positive and false positive predictions.

$$\text{precision} = \frac{\text{number of true positives}}{\text{number of true positives} + \text{number of false positives}}$$

Recall refers to the number of relevant predictions that are retrieved. In the context of our results, recall refers to the number of true positive predictions divided by the sum of the true positive and false negative predictions.

$$\text{recall} = \frac{\text{number of true positives}}{\text{number of true positives} + \text{number of false negatives}}$$

The f-score combines and balances both precision and recall and is referred to as the weighted average of both precision and recall.

$$\text{f-score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$





# User Interactions

## 3.1 Introduction

There are a number of user interactions the feature vector  $x$  can be comprised of. Each of these interactions types are compared to baselines to help gauge their effectiveness.

## 3.2 Interactions

For the interactions data type, the feature vectors are composed of the form  $x_i$  where  $i$  is an index into the vector and composed of the cross product of:

$$i = \{incoming, outgoing\} \times \{post, photo, video, link\} \times \{comment, tag, like\}$$

The alters of  $i$  can then be defined as all users who have interacted with the current user via some interaction  $i$ . The column is set to 1 if any of the alters defined by the current set  $i$  have also liked the item associated with the user, otherwise it is set to 0.

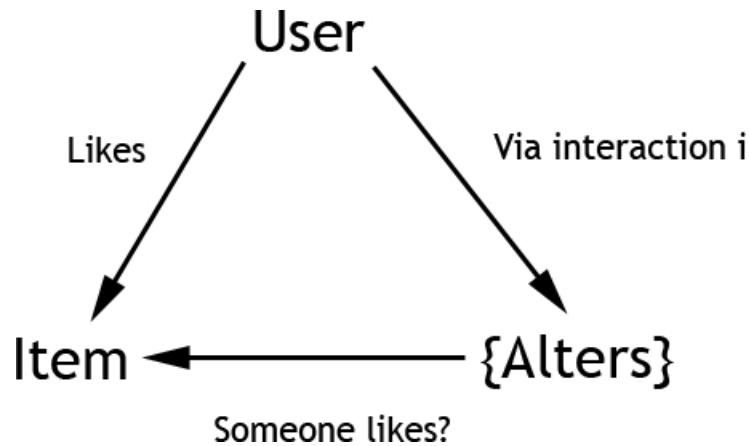
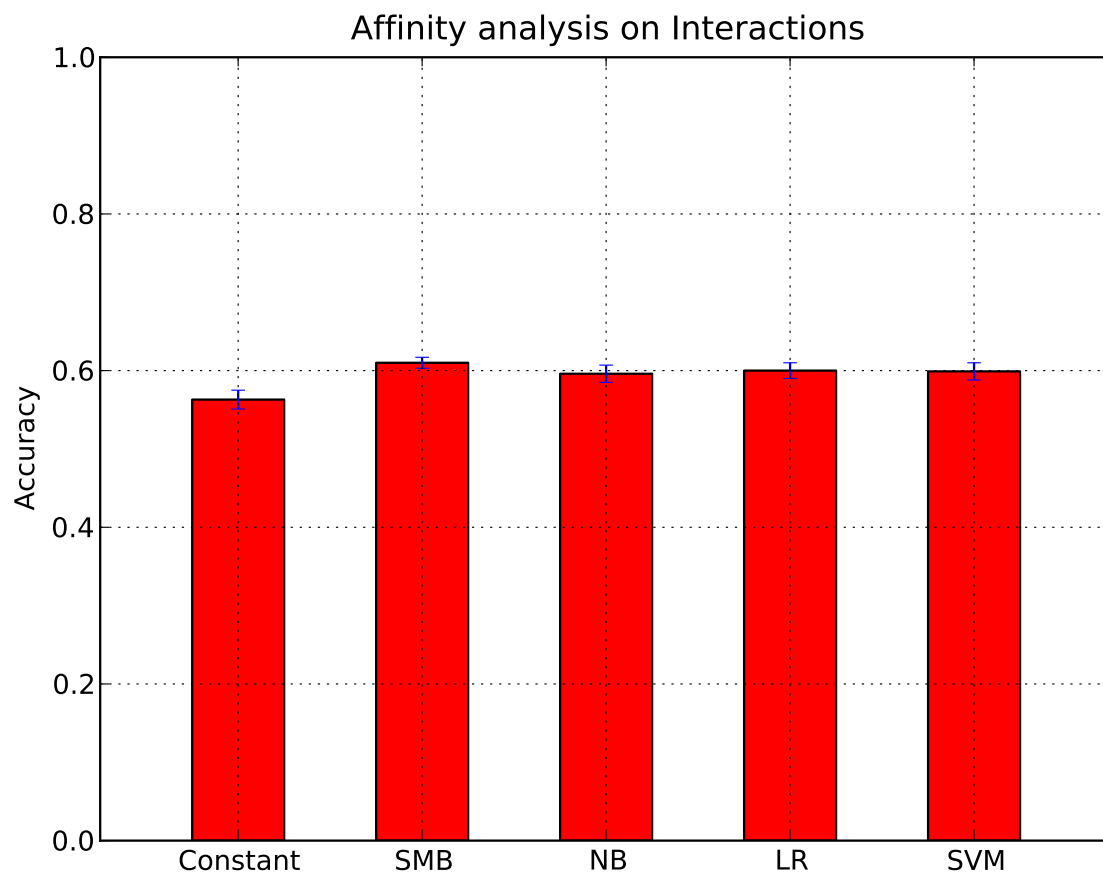
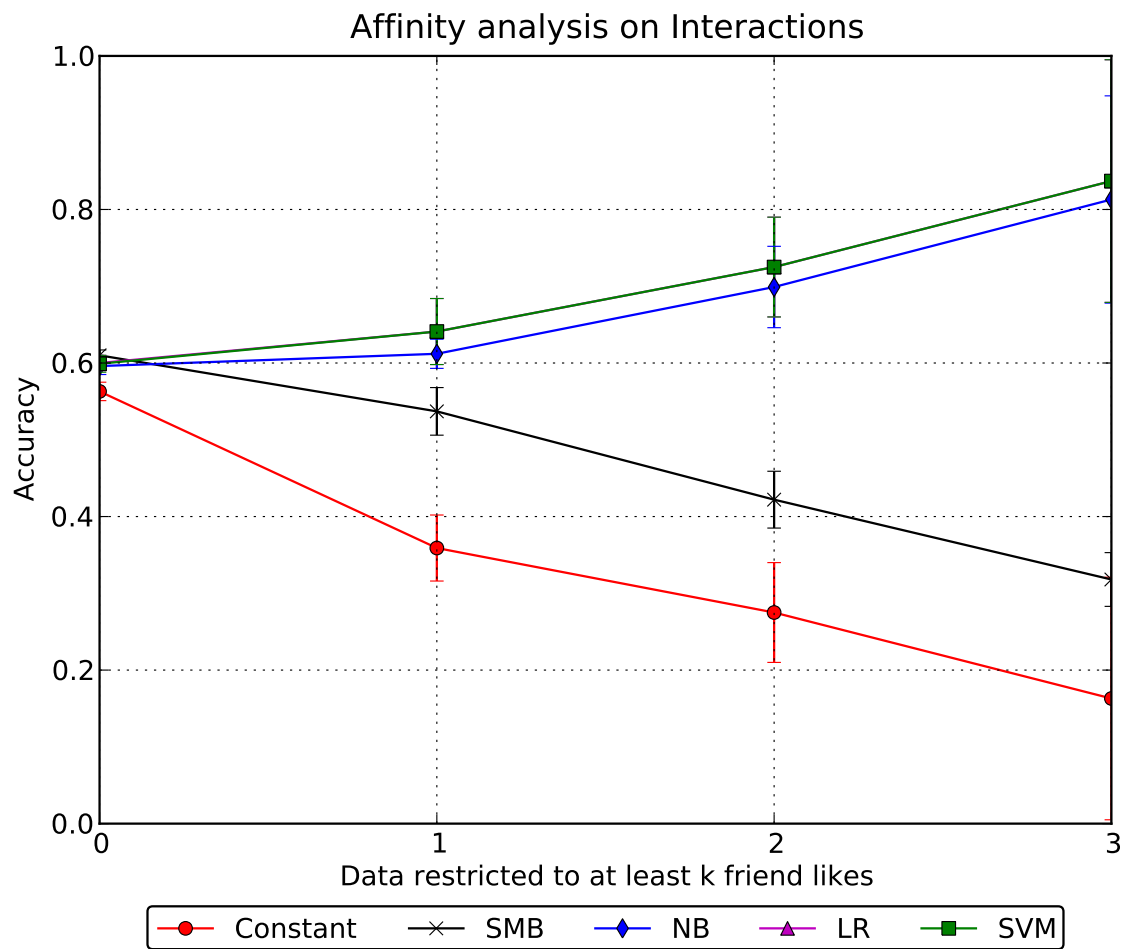


Figure 3.1: Predictors paradigm



**Figure 3.2:** Predictors paradigm

**Figure 3.3:** Predictors paradigm

### 3.3 Conversation

Given the nature of Facebook, it is possible for users to post messages to other users.

These messages can be broken down based on their directionality, either *outgoing* which are words sent to other users or *incoming* which are words received from other users.

The most commonly used words are below.

| Rank | Word     | Frequency |
|------|----------|-----------|
| 1    | :)       | 292,733   |
| 2    | like     | 198,289   |
| 3    | good     | 164,387   |
| 4    | thanks   | 159,238   |
| 5    | one      | 156,696   |
| 6    | love     | 139,939   |
| 7    | :p       | 121,904   |
| 8    | time     | 106,995   |
| 9    | think    | 106,459   |
| 10   | see      | 103,690   |
| 11   | nice     | 99,672    |
| 12   | now      | 94,947    |
| 13   | well     | 92,735    |
| 14   | happy    | 84,381    |
| 15   | :d       | 83,698    |
| 16   | much     | 78,719    |
| 17   | oh       | 77,321    |
| 18   | yeah     | 76,564    |
| 19   | back     | 76,032    |
| 20   | great    | 70,514    |
| 21   | going    | 70,447    |
| 22   | still    | 68,245    |
| 23   | new      | 67,430    |
| 24   | day      | 65,579    |
| 25   | come     | 63,837    |
| 26   | ;)       | 62,936    |
| 27   | year     | 61,771    |
| 28   | look     | 60,608    |
| 29   | yes      | 59,774    |
| 30   | want     | 59,514    |
| 31   | tag      | 58,633    |
| 32   | hahaha   | 57,448    |
| 33   | also     | 56,414    |
| 34   | need     | 55,921    |
| 35   | make     | 54,949    |
| 36   | sure     | 54,395    |
| 37   | thank    | 54,112    |
| 38   | people   | 53,211    |
| 39   | miss     | 53,182    |
| 40   | guys     | 52,855    |
| 41   | right    | 52,112    |
| 42   | best     | 51,941    |
| 43   | awesome  | 51,663    |
| 44   | hope     | 50,980    |
| 45   | 2        | 50,720    |
| 46   | next     | 50,375    |
| 47   | work     | 49,459    |
| 48   | way      | 49,358    |
| 49   | man      | 49,101    |
| 50   | :(       | 48,184    |
| 51   | j3       | 47,985    |
| 52   | even     | 47,480    |
| 53   | 4        | 46,068    |
| 54   | us       | 45,919    |
| 55   | pretty   | 44,804    |
| 56   | hey      | 44,614    |
| 57   | say      | 44,315    |
| 58   | better   | 43,357    |
| 59   | thanx    | 42,639    |
| 60   | bro      | 41,187    |
| 61   | take     | 41,081    |
| 62   | always   | 40,457    |
| 63   | wow      | 40,452    |
| 64   | pic      | 40,185    |
| 65   | though   | 40,032    |
| 66   | actually | 39,565    |
| 67   | last     | 39,175    |
| 68   | thats    | 38,833    |
| 69   | cool     | 37,844    |
| 70   | dear     | 37,328    |
| 71   | ok       | 36,441    |
| 72   | sorry    | 36,345    |
| 73   | never    | 36,000    |
| 74   | thing    | 35,941    |
| 75   | first    | 35,785    |
| 76   | looks    | 35,496    |
| 77   | night    | 35,475    |
| 78   | thought  | 34,458    |
| 79   | photo    | 33,989    |
| 80   | &        | 33,902    |

Table 3.1: Top conversation content data for all users

### 3.3.1 Outgoing

First we need to figure out the number of outgoing messages words are optimal for our classifiers.

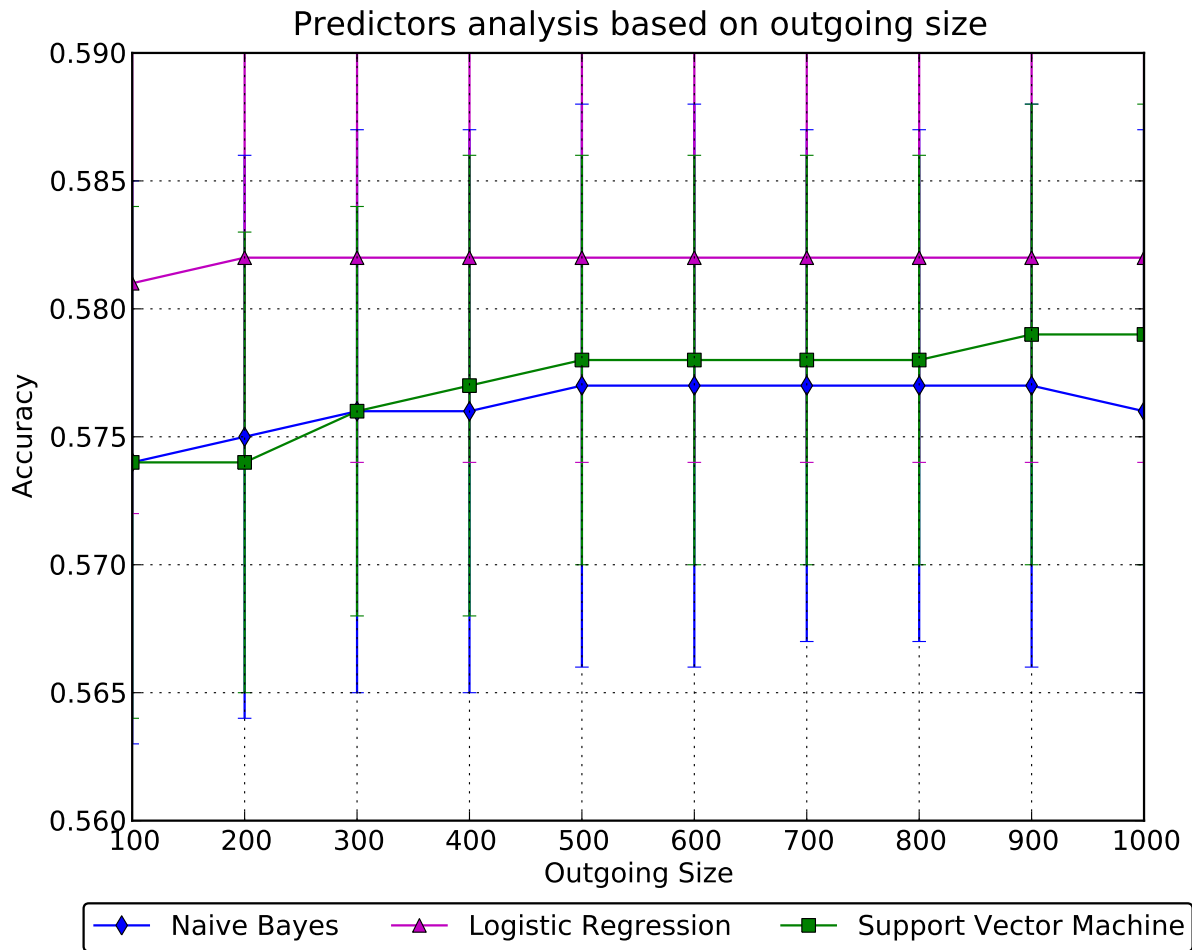


Figure 3.4: Predictors paradigm

### 3.3.2 Incoming

---

# User Preferences

---

## 4.1 Introduction

Facebook allows users to provide a vast array of personal traits and interests on their Facebook page.

Including:

- Demographics - age, gender, location, etc
- Group Memberships
- Personal Preferences - favourite books, favourite athletes, favourite sports, inspirational people, personal interests, etc
- Conversation Data - words sent, words received

In this section we will try to uncover which User Traits are indicative of item likes.

## 4.2 Demographics

Gender breakdown in the data set:

| Male | Female | Undisclosed |
|------|--------|-------------|
| 85   | 33     | 1           |

**Table 4.1:** Gender breakdown

There is a clear male bias in the data set.

Birthday breakdown in the data set:

| Year        | Frequency |
|-------------|-----------|
| Undisclosed | 1         |
| 1901-1905   | 1         |
| 1906-1910   | 0         |
| 1911-1915   | 1         |
| 1916-1920   | 0         |
| 1921-1925   | 0         |
| 1926-1930   | 0         |
| 1931-1935   | 0         |
| 1936-1940   | 1         |
| 1941-1945   | 0         |
| 1946-1950   | 0         |
| 1951-1955   | 0         |
| 1956-1960   | 2         |
| 1961-1965   | 1         |
| 1966-1970   | 4         |
| 1971-1975   | 10        |
| 1976-1980   | 12        |
| 1981-1985   | 25        |
| 1986-1990   | 34        |
| 1991-1995   | 25        |
| 1996-2000   | 2         |

**Table 4.2:** Birthday breakdown

Birthdays are grouped in a distinct range, most users in this data set are in the age range of 18 – 30.



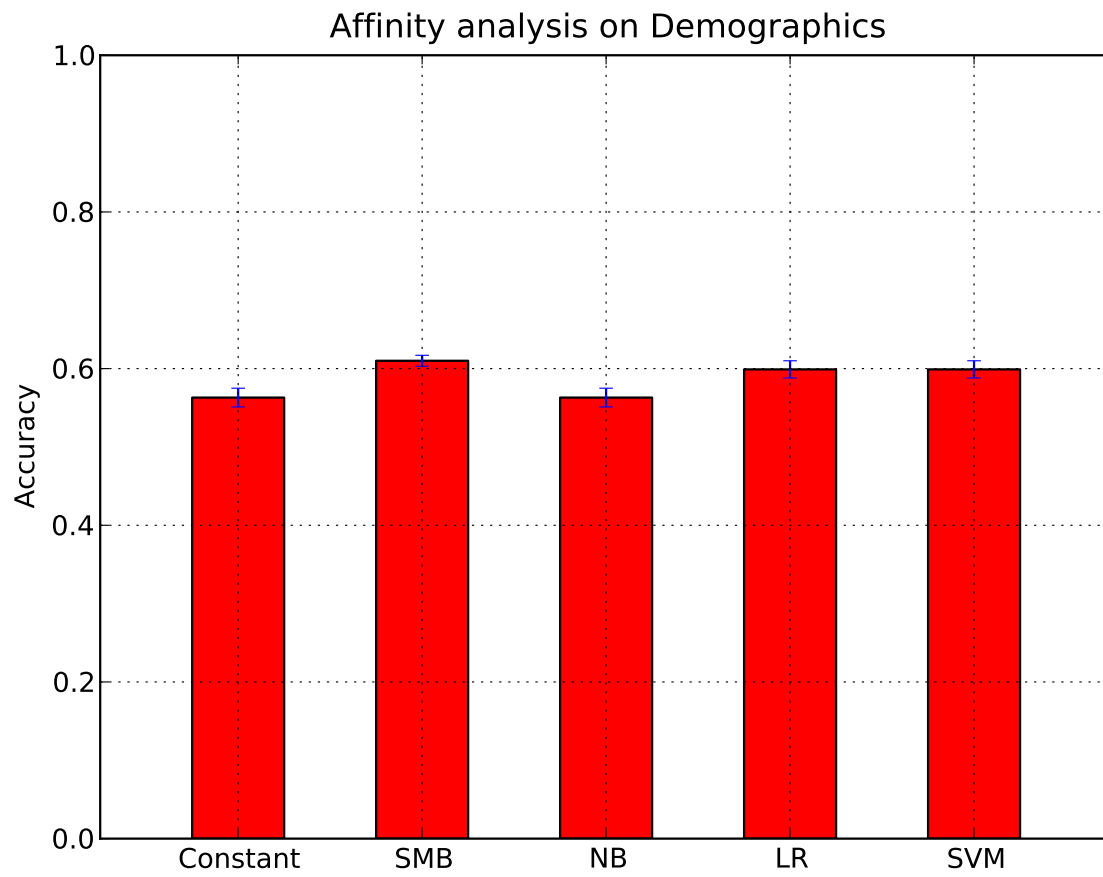
---

Location breakdown in the data set:

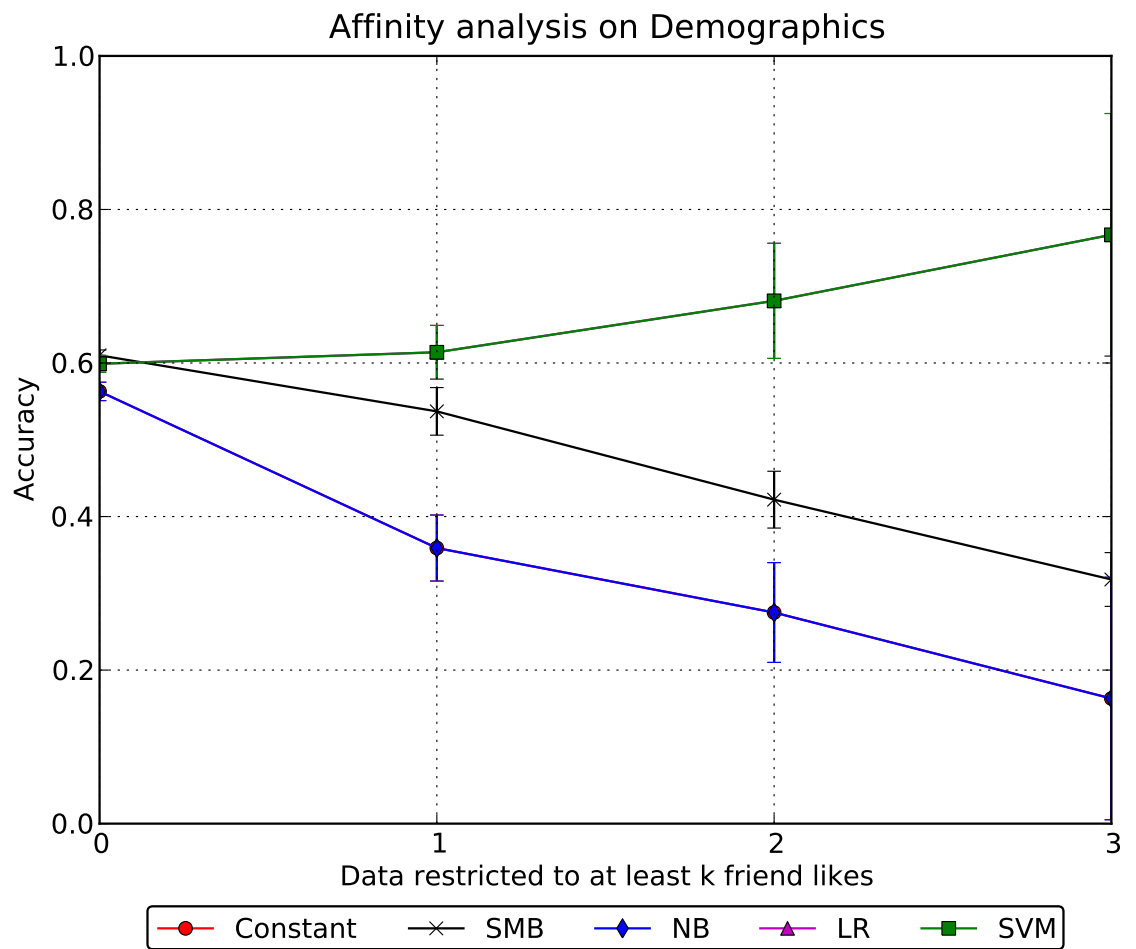
| Location   | Frequency |
|--|-----------|
| Undisclosed                                      | 33        |
| Ahmedabad, India                                 | 1         |
| Bangi, Malaysia                                  | 1         |
| Bathurst, New South Wales                        | 1         |
| Bellevue, Washington                             | 1         |
| Braddon, Australian Capital Territory, Australia | 1         |
| Brisbane, Queensland, Australia                  | 2         |
| Canberra, Australian Capital Territory           | 56        |
| Culver City, California                          | 1         |
| Frederick, Maryland                              | 3         |
| Geelong, Victoria                                | 1         |

**Table 4.3:** Location breakdown

Given the fact that most users are either situated in the ACT (location of the app development and deployment) or are undisclosed, location information in this data set will not be useful.



**Figure 4.1:** Predictors paradigm

**Figure 4.2:** Predictors paradigm

### 4.3 Traits

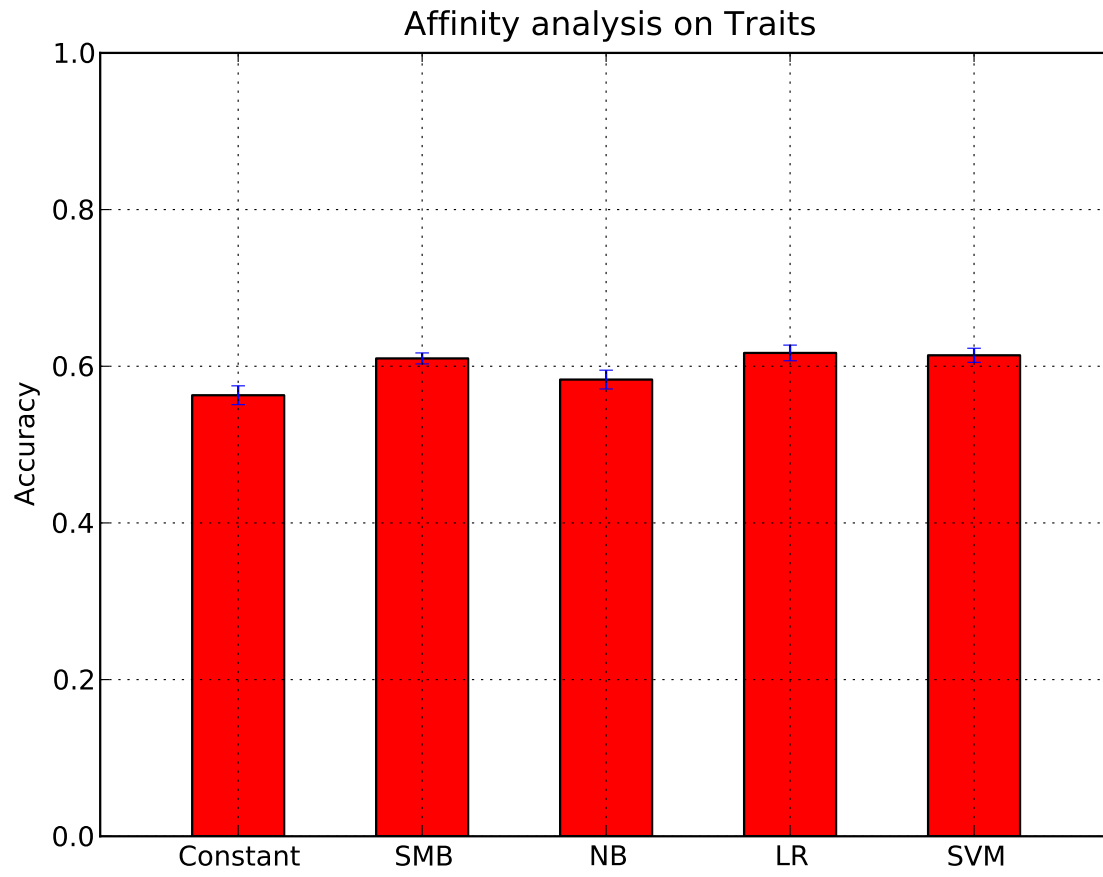
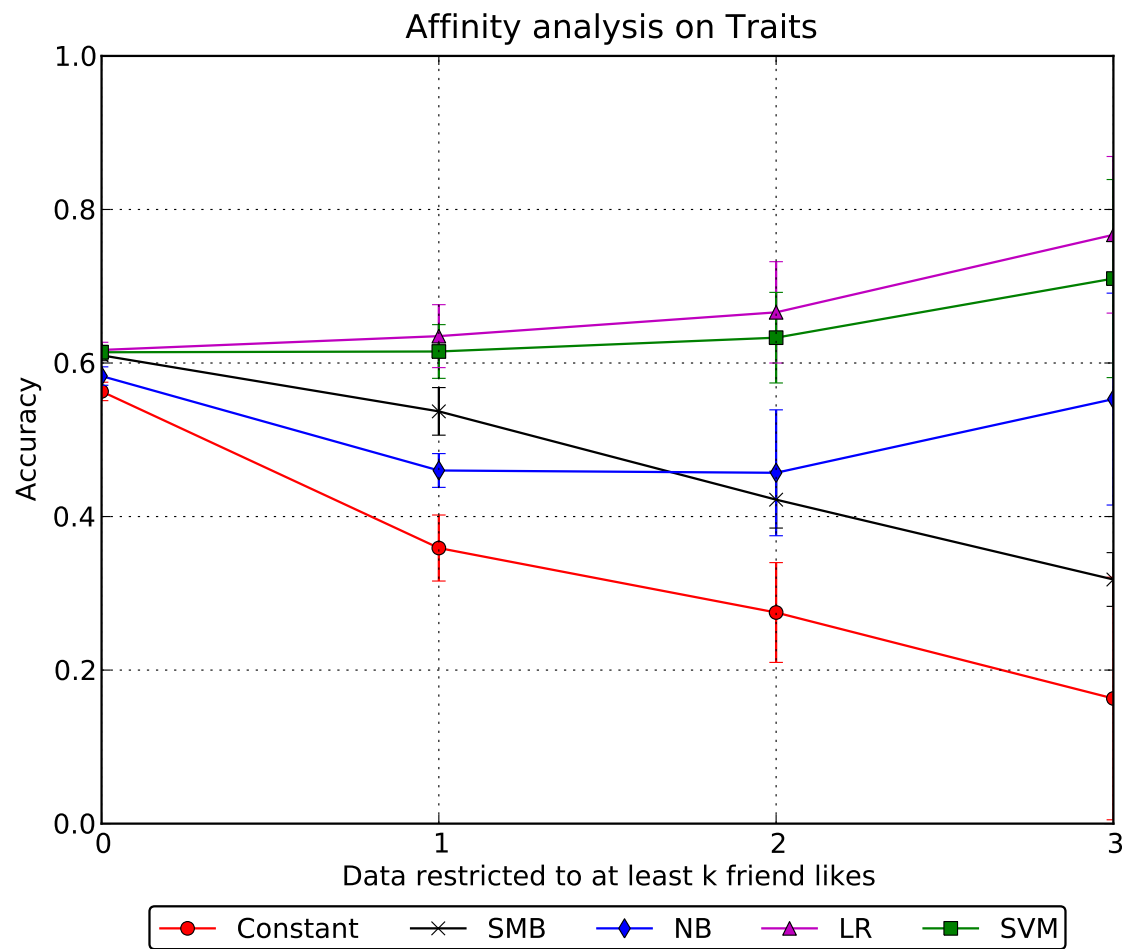


Figure 4.3: Predictors paradigm

**Figure 4.4:** Predictors paradigm

| Count | Activity    |
|-------|-------------|
| 10    | Sleeping    |
| 5     | Eating      |
| 5     | Reading     |
| 4     | Running     |
| 4     | Cycling     |
| 4     | Minecraft   |
| 4     | Programming |
| 3     | Android     |
| 3     | Cooking     |
| 3     | Video Games |
| 3     | Xbox 360    |
| 3     | Piano       |
| 3     | Guitar      |
| 3     | Badminton   |
| 3     | Chocolate   |

Table 4.4: Top activities for app users

| Count | Inspirational Person |
|-------|----------------------|
| 2     | Alan Turing          |
| 1     | Bender               |
| 1     | Maurice Moss         |
| 1     | Steve Jobs           |
| 1     | Sean Parker          |
| 1     | Pope Benedict XVI    |
| 1     | Martin Luther        |
| 1     | Alistair McGrath     |
| 1     | St Augustine         |
| 1     | Dennis Ritchie       |
| 1     | Linus Torvalds       |
| 1     | Richard Stallman     |
| 1     | C. S. Lewis          |
| 1     | Mike Oldfield        |
| 1     | Ryan Giggs           |

Table 4.5: Top inspirational people

| Count | Book   |
|-------|--|
| 7     | Harry Potter   |
| 4     | The Bible  |
| 3     | Harry Potter series  |
| 3     | Discworld  |
| 3     | That's 3 minutes of solid study, think I've earned 2hrs of Facebook time |
| 3     | Freakonomics   |
| 3     | Tomorrow when the War Began  |
| 2     | Magician   |
| 2     | Hitchhiker's Guide To The Galaxy   |
| 2     | The Discworld Series   |
| 2     | Terry Pratchett  |
| 2     | Terry Pratchett  |
| 2     | George Orwell  |
| 2     | Lord Of The Rings  |
| 2     | Goosebumps   |

Table 4.6: Top books for app users

---

| Count | Athlete   |
|-------|---|
| 4     | Roger Federer   |
| 4     | Rafael Nadal  |
| 3     | Maria Sharapova   |
| 2     | Leo Messi   |
| 1     | Andy Schleck  |
| 1     | Chrissie Wellington                                     |
| 1     | Emma Snowsill   |
| 1     | Emma Moffatt  |
| 1     | Brbara Riveros  |
| 1     | The Brownlee Brothers                                   |
| 1     | Marie Slamtoinette #1792                                |
| 1     | Wayne Rooney  |
| 1     | "you are what you eat" "I dont remember eating a Tank." |
| 1     | Nemanja Vidic   |
| 1     | Ryan Giggs  |

Table 4.7: Top athletes for app users

| Count | Interest                |
|-------|-------------------------|
| 5     | Movies                  |
| 5     | Music                   |
| 3     | Cooking                 |
| 3     | Sports                  |
| 2     | Psychology              |
| 2     | Internet                |
| 2     | Video Games             |
| 2     | Martial arts            |
| 2     | Literature              |
| 2     | Economics               |
| 2     | Tennis                  |
| 2     | Badminton               |
| 2     | Artificial intelligence |
| 2     | Computers               |
| 2     | Travel                  |

Table 4.8: Top interests for app users

| Count | Music           |
|-------|-----------------|
| 9     | Daft Punk       |
| 9     | Muse            |
| 8     | Michael Jackson |
| 8     | Pink Floyd      |
| 8     | Lady Gaga       |
| 7     | Linkin Park     |
| 7     | Avril Lavigne   |
| 6     | Radiohead       |
| 6     | Rihanna         |
| 6     | Coldplay        |
| 6     | Green Day       |
| 6     | Katy Perry      |
| 6     | Taylor Swift    |
| 5     | Gorillaz        |
| 5     | Queen           |

Table 4.9: Top music for app users

| Count | Movie  |
|-------|--|
| 9     | Inception  |
| 8     | Avatar   |
| 8     | Fight Club   |
| 7     | The Lord of the Rings Trilogy (Official Page)          |
| 6     | Star Wars  |
| 6     | I wouldnt steal a car, But i'd download one if i could |
| 6     | WALL-E   |
| 6     | Scott Pilgrim vs. the World                            |
| 6     | Toy Story  |
| 6     | Shrek  |
| 5     | Batman: The Dark Knight                                |
| 5     | Harry Potter   |
| 4     | The Matrix   |
| 4     | The Social Network Movie                               |
| 4     | Monsters, Inc.   |

Table 4.10: Top movies for app users

| Count | Sport             |
|-------|-------------------|
| 8     | Badminton         |
| 5     | Basketball        |
| 3     | Cycling           |
| 3     | Volleyball        |
| 2     | Starcraft II      |
| 2     | Football en salle |
| 2     | Swimming          |
| 2     | Towel Baseball    |
| 2     | Tennis            |
| 1     | Soccer            |
| 1     | Taekwondo         |
| 1     | Rock climbing     |
| 1     | In The Groove     |
| 1     | Darts             |
| 1     | Table tennis      |

Table 4.11: Top sports for app users

| Count | Television Show       |
|-------|-----------------------|
| 20    | The Big Bang Theory   |
| 19    | How I Met Your Mother |
| 14    | The Simpsons          |
| 13    | Top Gear              |
| 12    | Futurama              |
| 12    | Scrubs                |
| 11    | Black Books           |
| 10    | Black Books           |
| 10    | South Park            |
| 10    | Family Guy            |
| 9     | The Daily Show        |
| 8     | The IT Crowd          |
| 8     | FRIENDS (TV Show)     |
| 7     | True Blood            |
| 7     | MythBusters           |

Table 4.12: Top television shows for app users



---

| Count | Team                                       |
|-------|--|
| 5     | Manchester United                          |
| 2     | Bear Grylls cameraman appreciation society |
| 2     | Real Madrid C.F.                           |
| 2     | Liverpool FC                               |
| 1     | Leopard Trek                               |
| 1     | British Triathlon                          |
| 1     | TeamCWUK                                   |
| 1     | Surly Griffins                             |
| 1     | Canberra Raiders                           |
| 1     | Kolkata Knight Riders                      |
| 1     | Brisbane Roar FC                           |
| 1     | Brisbane Broncos                           |
| 1     | Cricket Australia                          |
| 1     | — Manchester United Fans —                 |
| 1     | Juventus                                   |

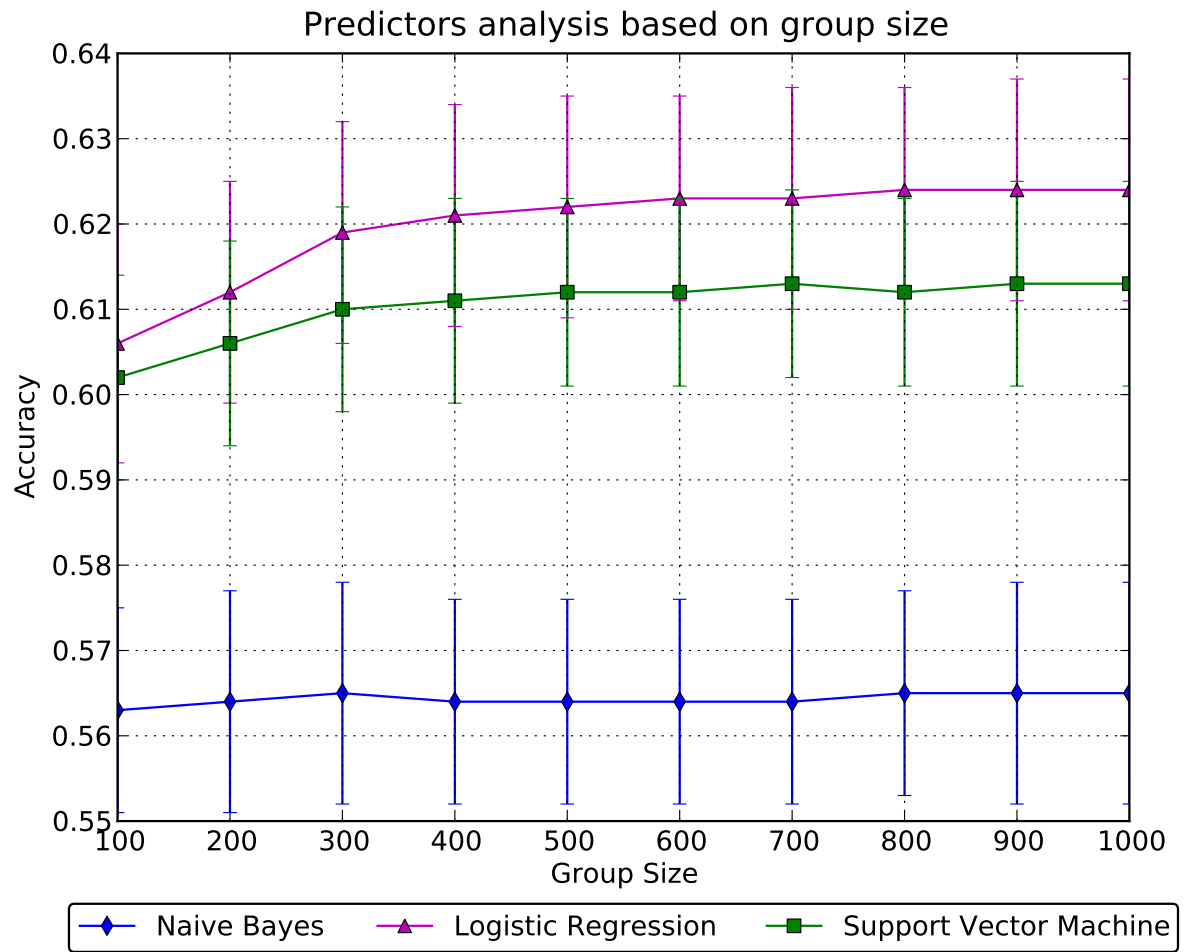
**Table 4.13:** Top teams for app users

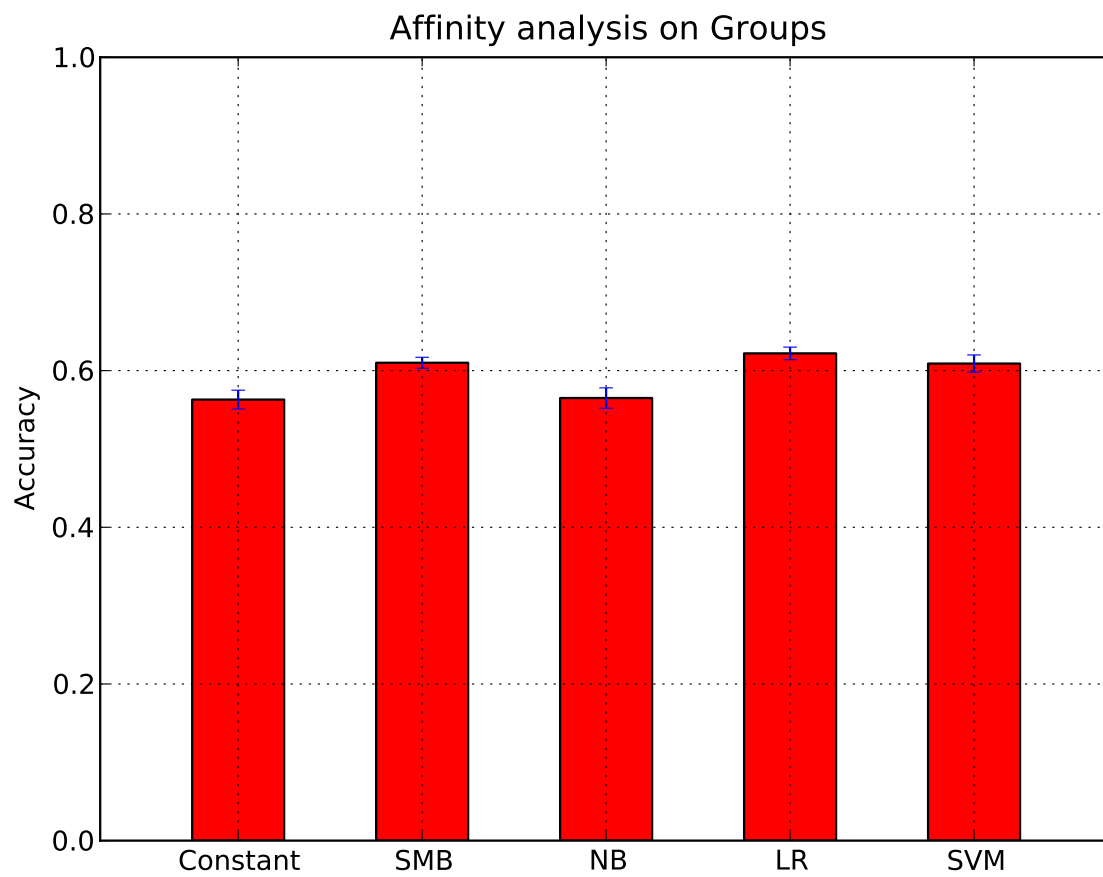
## 4.4 Groups

The most popular groups for app users are shown below.

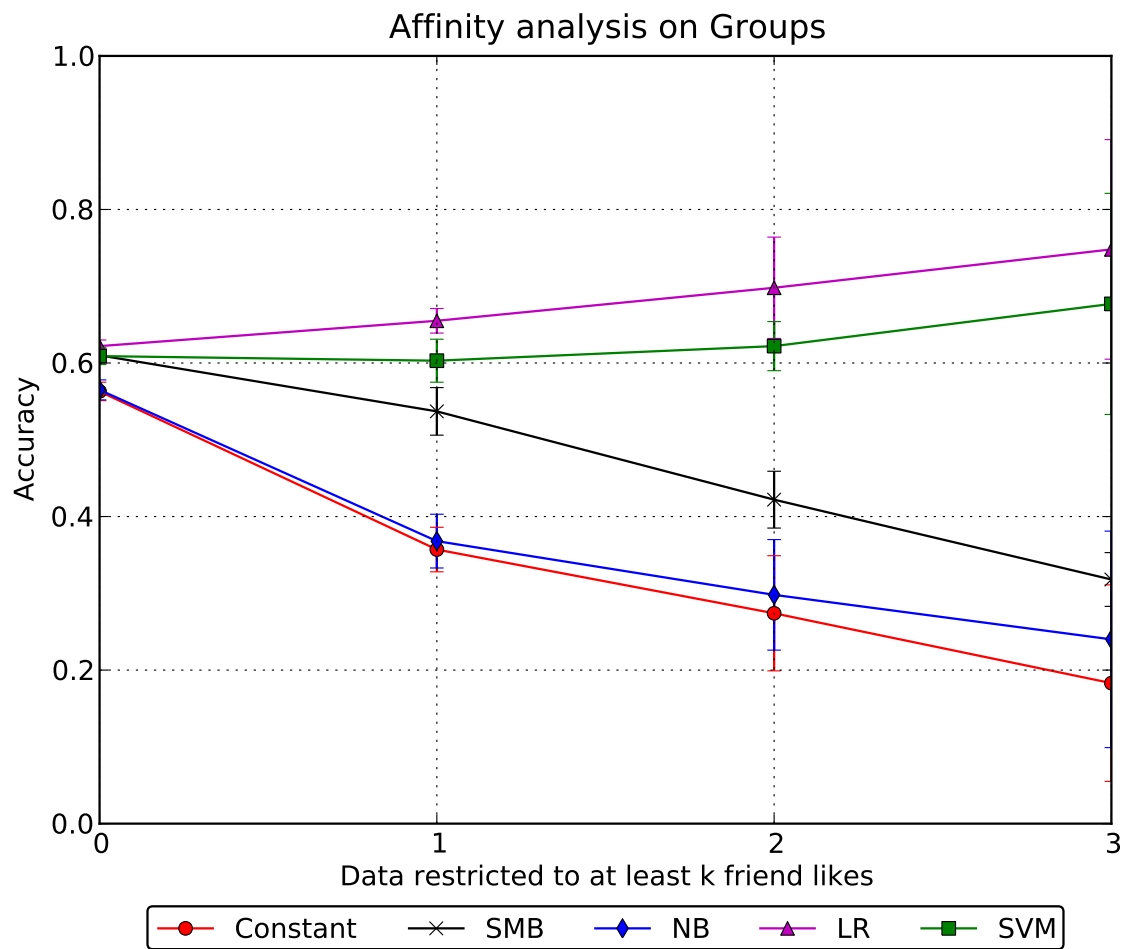
| Group Name | Frequency   |
|------------|---|
| 27         | ANU StalkerSpace  |
| 20         | Facebook Developers   |
| 15         | ANU CSSA  |
| 14         | CSSA  |
| 13         | Australian National University  |
| 11         | ANU - ML and AI Stanford Course   |
| 10         | iDiscount ANU   |
| 10         | Our Hero: Clem Baker-Finch  |
| 9          | Students In Canberra  |
| 7          | I grew up in Australia in the 90s   |
| 7          | Grow up Australia - R18+ Rating for Computer Games                          |
| 7          | ANU Engineering Students' Association (ANUESA) 2010                         |
| 7          | ANU Postgraduate and Research Student Association (PARSA)                   |
| 6          | No, I Don't Care If I Die At 12AM, I Refuse To Pass On Your Chain Letter.   |
| 6          | No Australian Internet Censorship   |
| 6          | The Chaser Appreciation Society   |
| 6          | Feed a Child with a Click   |
| 6          | ANU Mathematics Society   |
| 6          | ANU International Student Services, CRICOS Provider Number 00120C           |
| 6          | 2011 New & Returning Burton & Garran Hall                                   |
| 5          | If You Can't Differentiate Between "Your" and "You're" You Deserve To Die   |
| 5          | Keep the ANU Supermarket!!!   |
| 5          | If 1m people join, girlfriend will let me turn our house into a pirate ship |
| 5          | The Great Australian Internet Blackout                                      |
| 5          | When I was your age, Pluto was a planet.                                    |
| 5          | Australian National University  |
| 5          | ANU International Students' Department                                      |
| 5          | We Won't Accept It - No To Mandatory Internet Censorship In Australia       |
| 5          | HvZ VS Sprinklers   |
| 5          | SC2 in Canberra   |
| 5          | An Arbitrary Number of People Demanding That Some Sort Of Action Be Taken   |
| 5          | PETITION FOR FACEBOOK TO INSTALL A DISLIKE BUTTON - the original            |

**Table 4.14:** App users groups breakdown for range 5+

**Figure 4.5:** Predictors paradigm



**Figure 4.6:** Predictors paradigm

**Figure 4.7:** Predictors paradigm

## 4.5 Pages

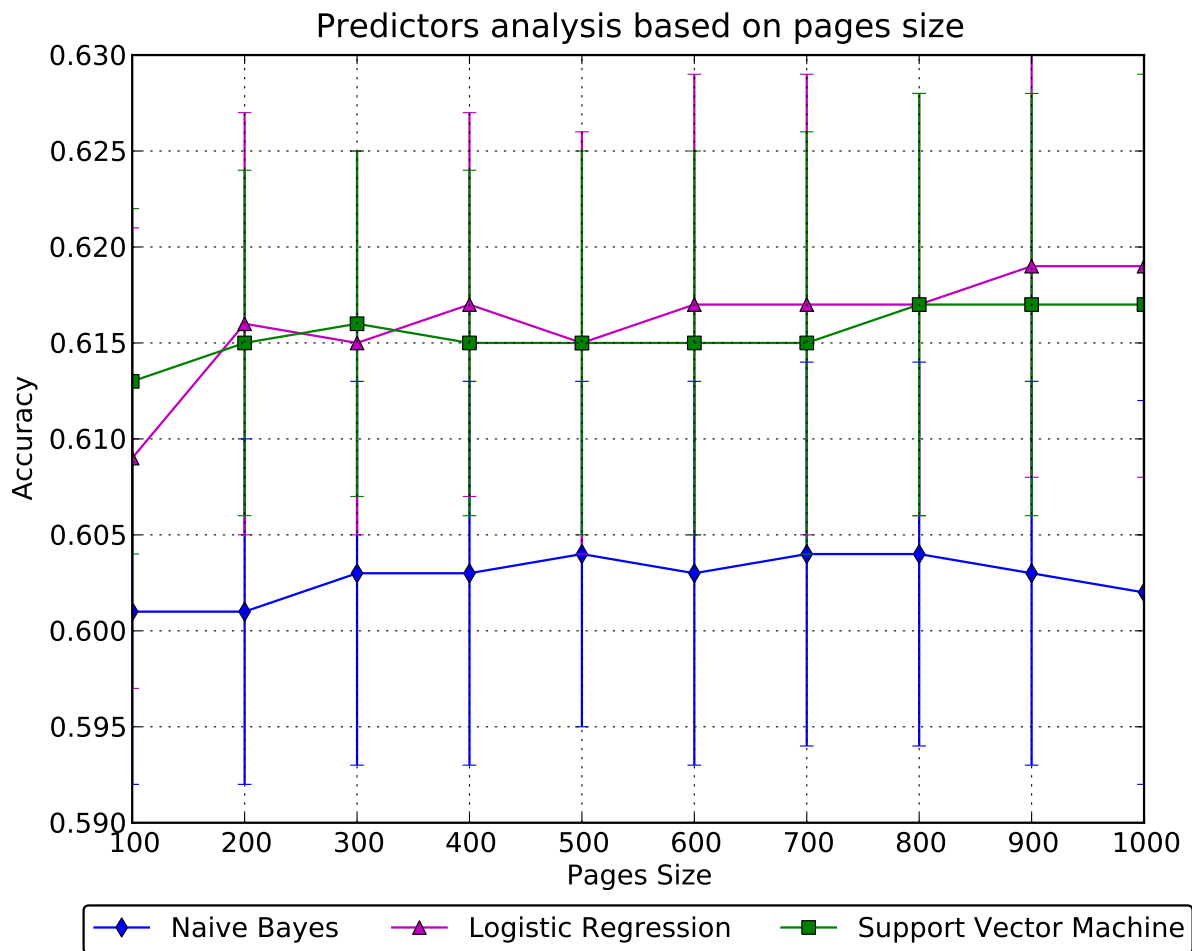
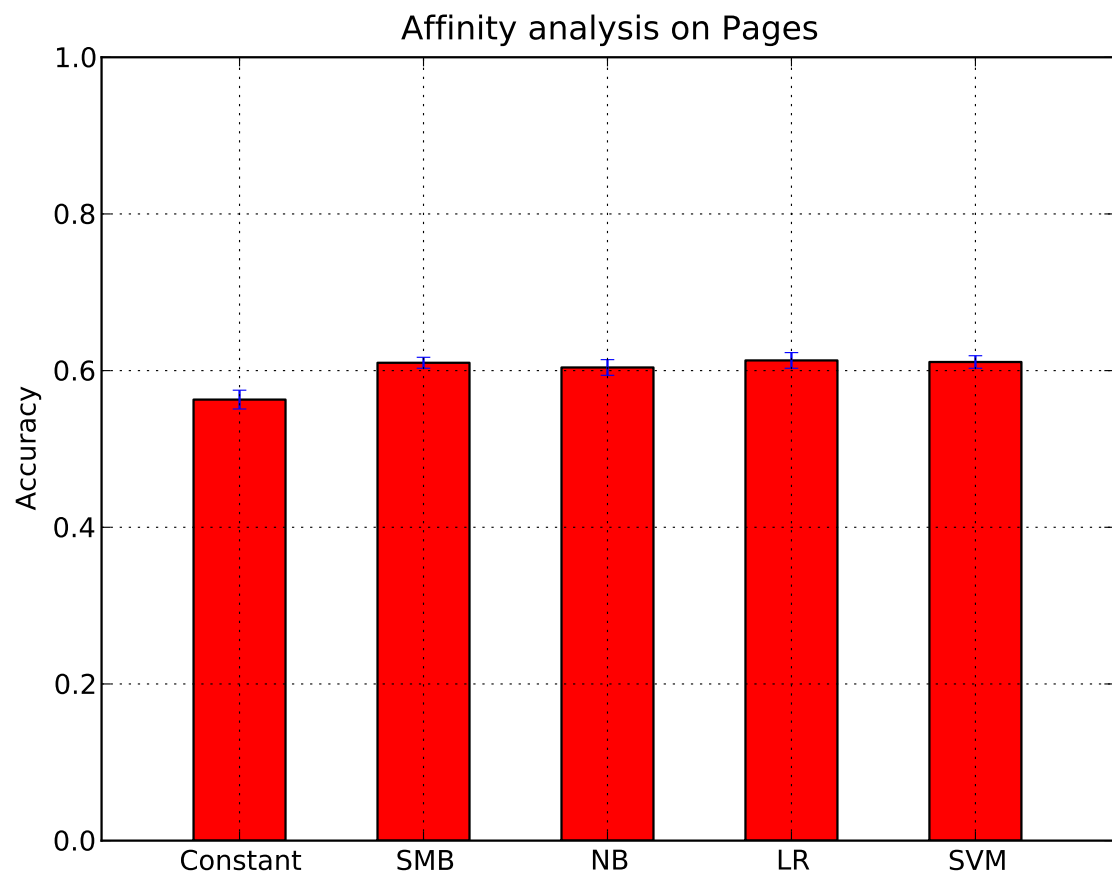


Figure 4.8: Predictors paradigm



**Figure 4.9:** Predictors paradigm

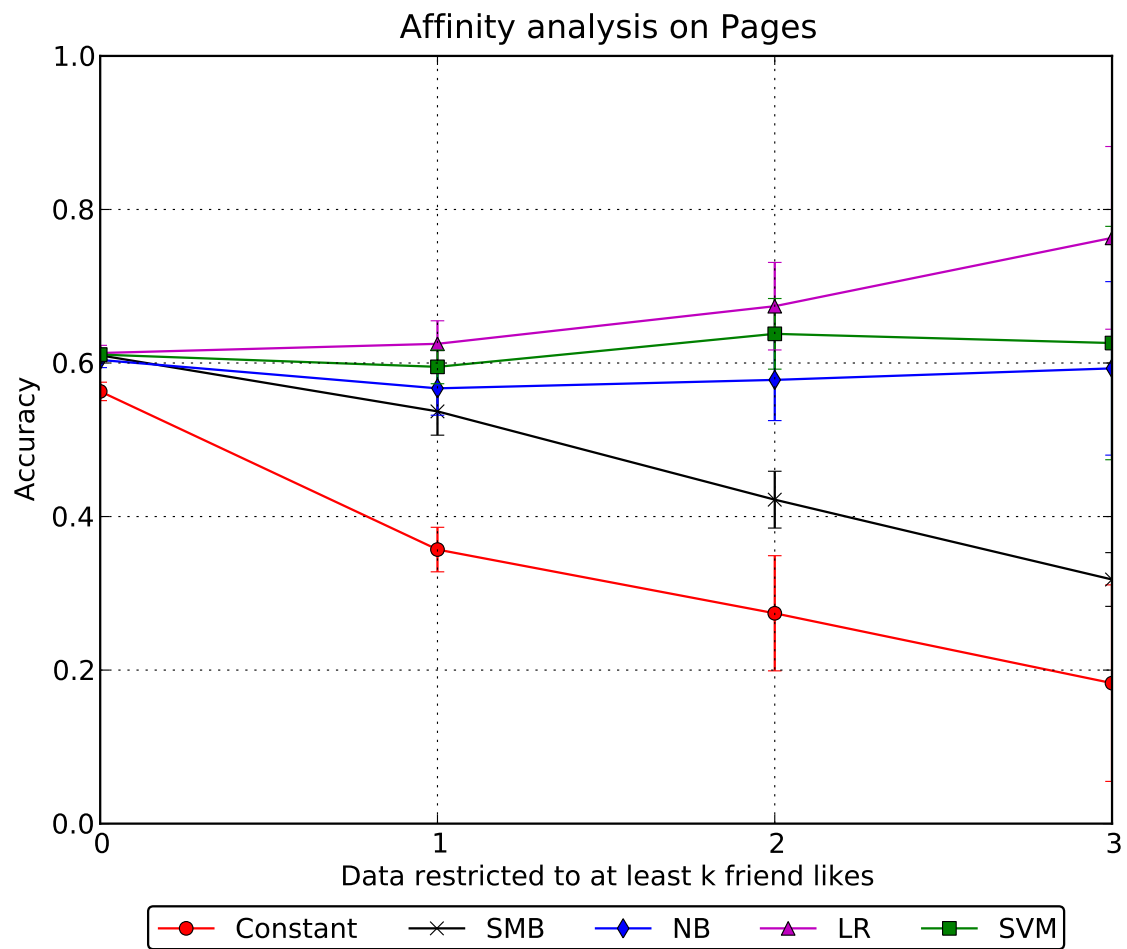


Figure 4.10: Predictors paradigm



---

# Bayesian Model Averaging

---

## 5.1 Introduction

## 5.2 Derivation

## 5.3 Results



---

# Conclusions

---

## 6.1 Summary

## 6.2 Future Work

