PREDICTION USING R

By

Waqas Ahmed

Independent Study
Submitted to the

Faculty of Computing & Engineering Sciences
Shaheed Zulfikar Ali Bhutto Institute of Science and Technology
In partial fulfilment of the requirements
Of the degree of
Master of Science

July, 2015

ACKNOWLEDGEMENTS

Thanks to Almighty who blessed me with skills and guided me to complete this Independent study research. Also I would like to express my sincere gratitude to my supervisor Prof. Dr. Tariq Rahim Soomro who provided me with the aspiring guidance and valuable advice regarding my area of studies for this Master's degree. I also appreciate all the suggestions and constructive criticism made by my friends and family on issues relating to my topic.

I acknowledge that whatever good is in this study is by the grace of Almighty and all mistakes are due to me, the servant of Almighty.

Waqas Ahmed

147301

ABSTRACT

One of the most exciting areas in all of data science right now is wearable. Companies like Fitbit, Nike, and Jawbone Up are racing to develop the most advanced algorithms to attract new users. In this study, analysis on two datasets has been done, which are 1) Titanic: learning from natural disaster, 2) Smartphone user activity prediction. In section 1, a brief introduction about data science is discussed, followed by section 2 where materials and methods described that were used to analyze and make precise predictions. In section 3, results are shown which were obtained from the participated kaggle competition. In the last section, conclusion and the room for improvement is briefly mentioned.

TABLE OF CONTENTS

TABLE OF FIGURES

CHAPTER 1

1. INTRODUCTION

1.1 About Data Science

Generally, Data Science is learning from data by the extraction of knowledge from it. [1][2] Data scientists today are people who have a blend of many different skills. The Venn diagram in Figure 1.1 shows a definition of a data scientist beautifully.

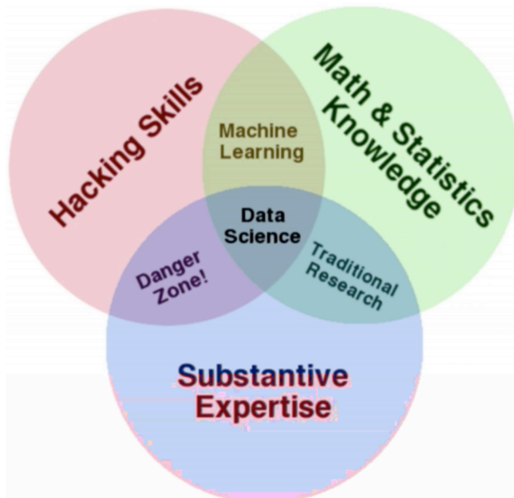


Figure 1.1: Data Scientist skills

A data scientist is the one who possess mathematics and statistics skills, which allows them to identify interesting insights in a sea of data. They also have the programming skills to code up statistical models and get data from a variety of different data sources. Furthermore, a data scientist is someone who knows how to ask the right questions and translate those questions into a sound analysis. After doing the analysis, they have the communication skills to report their findings in a way that people can easily understand. In other words, data scientists have the ability to perform complicated analysis on huge data sets. Once they've done this, they also have the ability to write and make informative graphs to communicate their findings to others. Also, a data scientist is a team player where other members have similar or more knowledge and skills that improves his or hers [3].

Looking at data science from the business angle, it is an integral part of competitive intelligence, a recently developing field that envelops various exercises, such as data mining and data analysis [4], as shown in Figure 1.2.

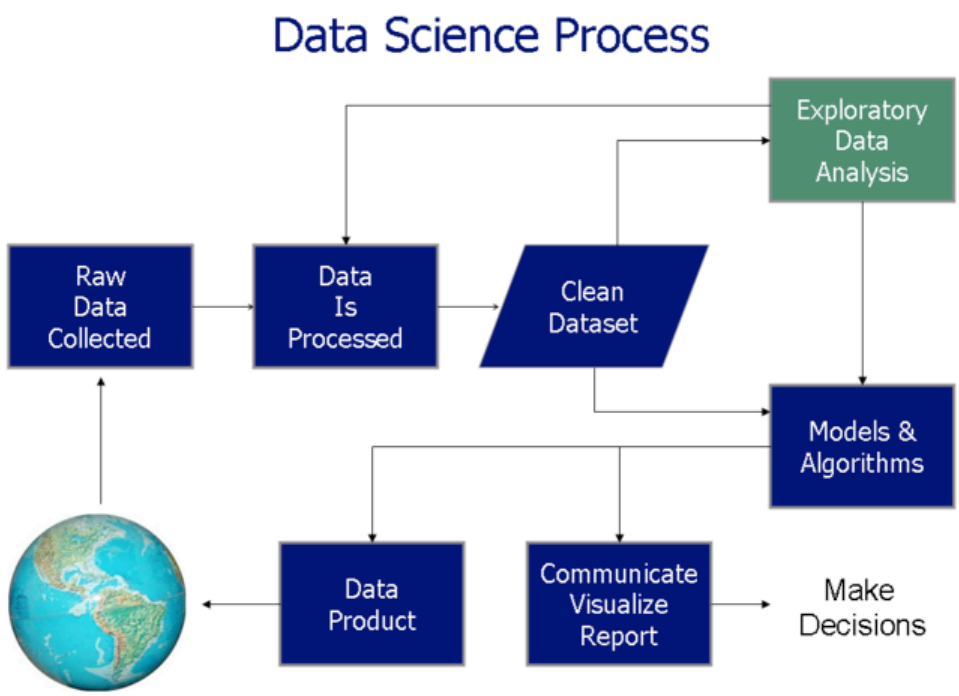


Figure 1.2: Data Science Process

Here are some things that a data scientist may do in his or her daily work as shown in Figure 1.3. They might wrangle data. That is:

- collect data from the real world
- process the collected data, and
- clean and tidy the data set such that it can be analyzed.
- After having tidy data set, trends may be analyzed in the existing data or they try to make data driven predictions about the future using the data at hand.

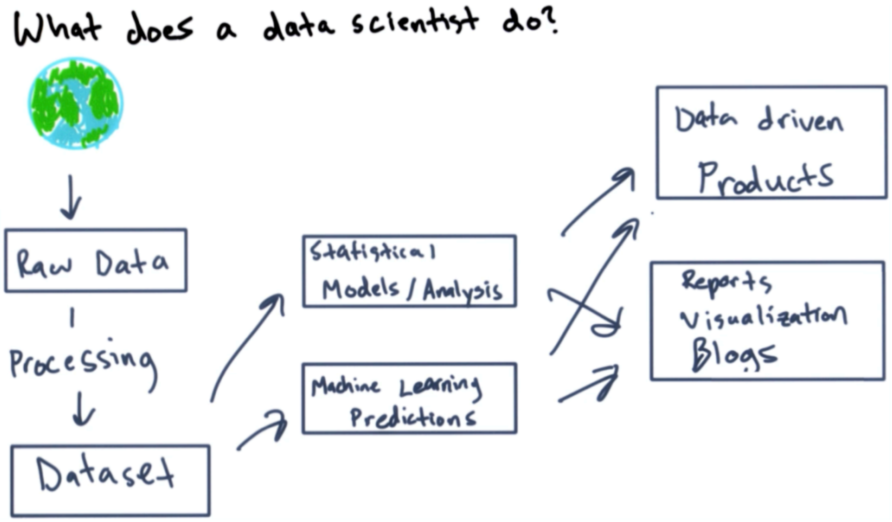


Figure 1.3: What does Data Scientist do

On the basis of these models or predictions, building data driven products is possible moreover, data scientists can communicate their findings with others, data scientists and the general public alike. File visualizations, charts, blog posts or reports.

1.2 Topic Exploration & Literature reviews

One of the most exciting areas in all of data science right now is wearable computing - according to this article [5]. Companies like Fitbit, Nike, and Jawbone Up are racing to develop the most advanced algorithms to attract new users. The dataset to be used to demonstrate independent study represents data collected from the accelerometers from the Samsung Galaxy S smartphone. A full description is available at the site where the data was obtained [6].

1.3 Literature Review

Human Activity Recognition (HAR) is playing a central role in extending the range of possibilities of human-computer interaction by offering the input for the development of more interactive and cognitive environments [7]. A compilation of the recent proposals in the Human Activity and Motion Disorder Recognition were exposed accompanied by the results of the contest calling for innovative approaches to recognize activities of daily living (ADL) from a recently published data set [8].

Moreover, various systems, theories, models etc.[9][10] have been presented for human physical Activity Recognition (AR or HAR) using smartphone inertial sensors, mainly when these sensors are attached to the subject's body, since they allow continuous monitoring of numerous physiological signals such as walking, running, sitting, lying etc.

CHAPTER 2

2. MATERIAL AND METHODS

This Independent study was divided into weekly studies and meanwhile carry on with the analysis of the dataset for smartphone human activity prediction or recognition. Below is the breakdown of how the whole study was scheduled:

2.1 Learning Sources

Learning the "Intro to Data Science" (The Data Scientist's Toolbox) included the following:

- Introduction to basic tools
  - R
  - Rstudio
  - Git
  - Github
- Types of data questions
- Steps in a data analysis
- Putting the science in data science

Learning "R programming" included the following:

- Overview of R, R data types and objects, reading and writing data
- Control structures, functions, scoping rules, dates and times
- Loop functions, debugging tools
- Simulation, code profiling

Learning "Getting and Cleaning Data" (Data Analysis with R) included the following:

- Data collection
  - Raw files (.csv,.xlsx)
  - Databases (mySQL)
  - APIs
- Data formats
  - Flat files (.csv,.txt)

    - XML
    - JSON
- Making data tidy
- Distributing data
- Scripting for data cleaning

Learning “Exploratory Data Analysis” included the following:

- summarizing data techniques
- complex statistical models
- techniques to eliminate or sharpen potential hypotheses using data to express the world
- how to plot using R, also fundamental rules of constructing data graphics and more.

2.2 Datasets

2.2.1 Kaggle Competition - Titanic: Machine Learning from Disaster (Predicting Survival Rate of Passengers) [7]

2.2.2 Kaggle Competition - Human Activity Recognition Using Smartphones (Predicting user activity from smartphone accelerometer and gyroscope data) [8]

2.3 Methods

For the independent study, exploratory data analysis was conducted and R scripts file were created that explores the variables, structure, patterns, oddities, and underlying relationships of the chosen data sets.

The analysis was almost like a stream-of-consciousness as data analyst ask questions, create visualizations, and explore the data.

The purpose of this project was to demonstrate the ability to collect, work with, and clean a data set. The goal was to learn how to explore data and predict the outcome using machine learning.

2.3.1. Data Collection

The data consisted of accelerometer and gyroscope measurements from a Samsung Galaxy IIS phone taken from a sample of 30 test subjects. Each record consisted of 561 time-and frequency-domain variables and the activity that the user was engaged in at the time of measurement, which may be: “standing”, “sitting”, “laying”, “walk”, “walk downwards”, “walk upwards.”

2.3.2. Data Subsets

The data were split into two subsets–TRAINING, TEST–based on subject IDs.

2.3.3. Exploratory Analysis

Applied Machine Learning using Random Forest Analysis

Exploratory analysis was used to

(1) Identify any missing values and

(2) Identify biases in the original dataset and each of the subsets, and

(3) Determine any necessary transformations to be made on the data.

There were no missing values left in the dataset after this analysis.

2.3.4. Decision Trees

First all the slicing was done and then efforts were made to find subsets that have a higher chance of surviving. A decision tree automates this process, and outputs a flowchart-like structure that is easy to interpret.

Conceptually, the decision tree algorithm starts with all the data at the root node and scans all the variables for the best one to split on. Once a variable is chosen, split is done and the next step is to go down one level (or one node) and repeat. The final nodes at the bottom of the decision tree are known as terminal nodes, and the majority vote of the observations in that bucket determine how to predict for new observations that end up in that terminal node.

To create the first decision tree, R's rpart package was used. R packages are a collection of functions, data and compiled code that make the life of data analyst easier. Namely, instead of needing to write the algorithm from scratch just the ‘rpart’ R package can be used and its included decision tree algorithm, as shown in Figure 2.1.

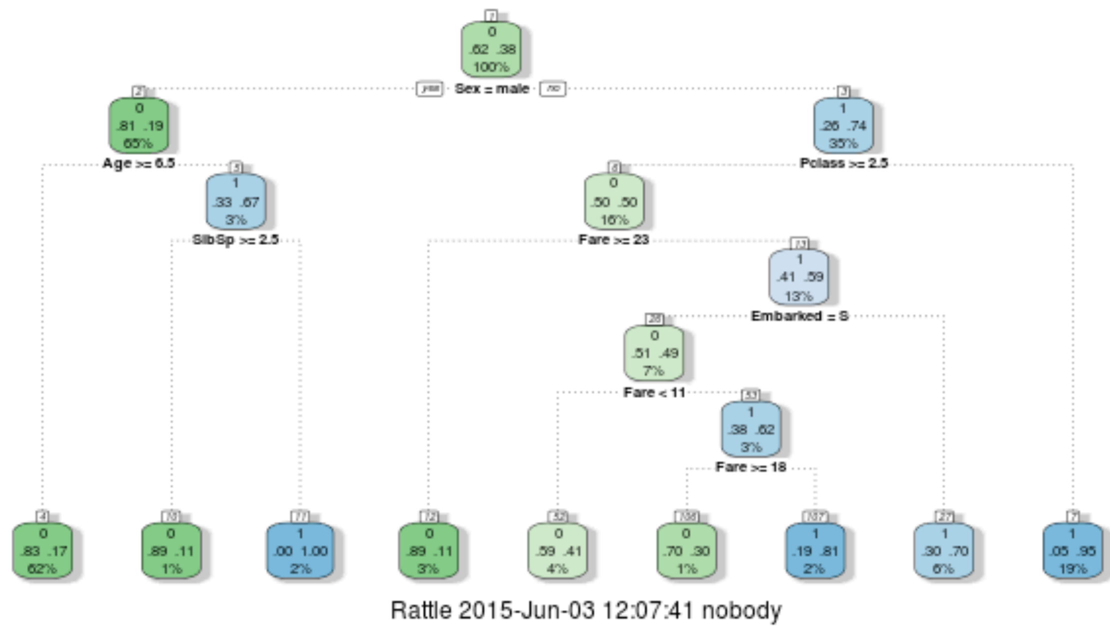


Figure 2.1: Decision tree of titanic dataset

2.3.5. What is a Random Forest Analysis

A brief study of Random Forests was done. However, since it's an often used machine learning technique, a general understanding and illustration of how to apply the technique using R was studied as well.

In layman terms, the Random Forest technique handles the over-fitting problem data analyst faces with decision trees. It grows multiple (very deep) classification trees using the training set. At the time of prediction, each tree is used to come up with a prediction and every outcome is counted as a vote. For example, if it has been trained 3 trees with 2 saying a passenger in the test set will survive and 1 says he will not, the passenger will be classified as a survivor. This approach of overtraining trees, but having the majority's vote count as the actual classification decision, avoids over-fitting, as show in Figure 2.2.

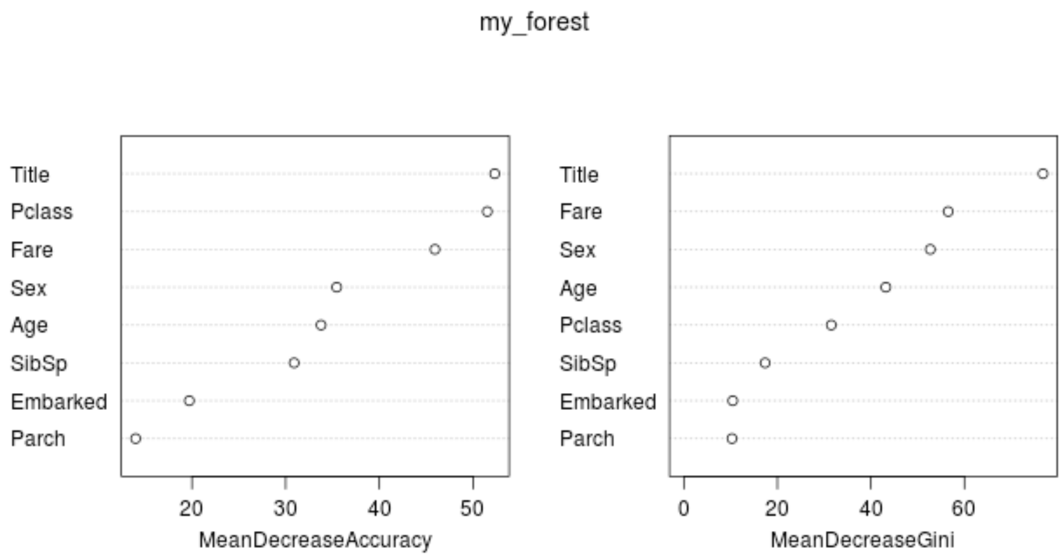


Figure 2.2:Importance of various factors in predicting the survival rate of passengers by applying random forest

2.3.6. Applied Machine Learning using Random Forest Analysis

After exploratory analysis, machine learning was introduced on the training datasets to predict required parameter, in this case survival rate of passengers and smartphone user activity.

CHAPTER 3

3. RESULT AND FINDINGS

Classification tree was made on the dataset of titanic disaster to find out the survival rate of the passengers onboard, which resulted in 80% correctness. Random forest analysis on the same dataset didn't improve the prediction precision. Moreover, participation in kaggle competition for the same ranked the model at 455th among 2972 participants [11], as shown in Figure 3.1 to 3.3.

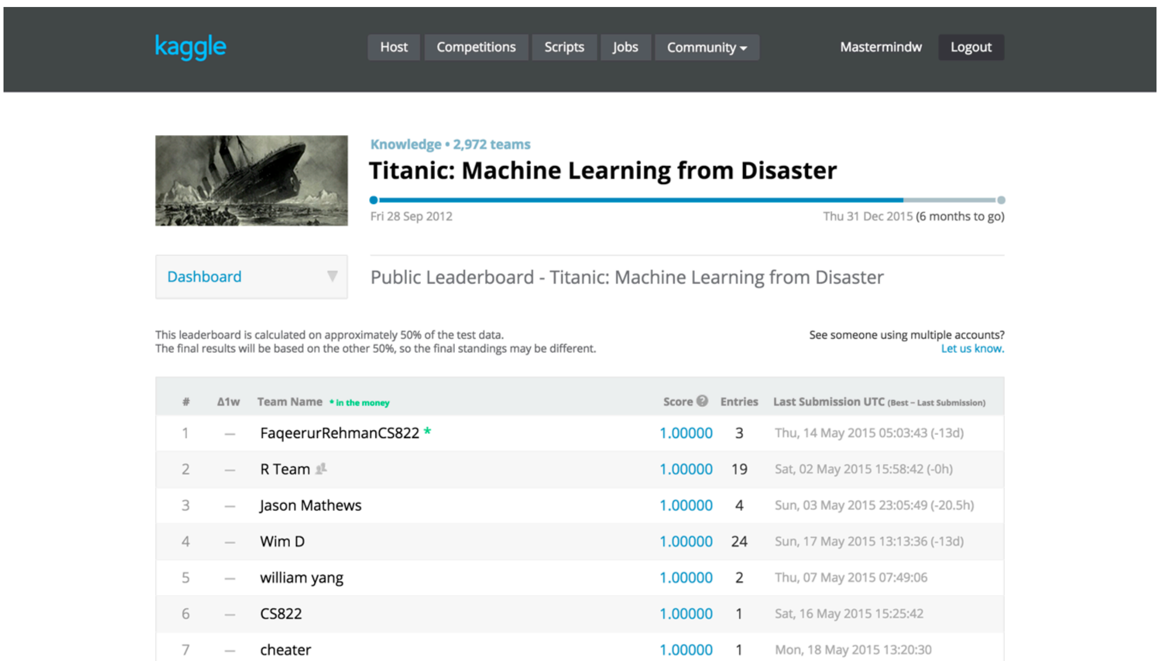


Figure 3.1: Snapshot of kaggle competition "Titanic: Machine Learning from Disaster" leaderboard initial screen

448	↓38	AmanPahariya	0.80383	4	Wed, 03 Jun 2015 07:38:24 (-21.9h)
449	↓38	Kaspar	0.80383	4	Wed, 03 Jun 2015 09:45:02 (-21.1h)
450	↓38	miziacz	0.80383	3	Tue, 02 Jun 2015 13:45:02 (-0.3h)
451	↓38	Stanley 2	0.80383	4	Wed, 03 Jun 2015 21:30:57 (-25.5h)
452	↓38	Cyber_Julia	0.80383	3	Tue, 02 Jun 2015 21:57:10
453	↓38	Pallen	0.80383	4	Wed, 03 Jun 2015 05:41:39 (-0.2h)
454	↓38	newbie 2	0.80383	8	Mon, 08 Jun 2015 10:25:37 (-5d)
455	**↓38**	**Mastermindw**	**0.80383**	**5**	**Thu, 25 Jun 2015 20:37:40 (-22.1d)**

Your Best Entry ↑
Your submission scored **0.77512**, which is not an improvement of your best score. Keep trying!

456	↓38	RachEl Lee	0.80383	3	Wed, 03 Jun 2015 20:26:43 (-0.2h)
457	↓38	Greg Lam	0.80383	2	Wed, 03 Jun 2015 22:37:17 (-0.2h)
458	↓38	thrivix	0.80383	4	Thu, 04 Jun 2015 00:02:43 (-0.9h)
459	↓38	Shannon Fenn	0.80383	3	Thu, 04 Jun 2015 04:03:45 (-0.2h)
460	↓38	McCartney Taylor	0.80383	5	Fri, 05 Jun 2015 01:15:52 (-19.2h)
461	↓38	Ishu Goyal	0.80383	4	Fri, 05 Jun 2015 06:50:45 (-20.6h)
462	↓38	Helani	0.80383	4	Fri, 05 Jun 2015 07:02:02 (-20h)
463	↓38	Neethu M V	0.80383	3	Thu, 04 Jun 2015 14:05:34 (-0.2h)

Figure 3.2: Snapshot of kaggle competition "Titanic: Machine Learning from Disaster" leaderboard highlighted rank screen

2957	↓357	YunSangLee	0.54067	1	Sat, 06 Jun 2015 08:06:44
2958	↓357	Alex_Royan	0.53589	2	Sat, 13 Jun 2015 21:23:29 (-0.5h)
2959	↓357	Manila Heat	0.52153	1	Tue, 09 Jun 2015 10:38:58
2960	↓357	pauljoseph	0.50718	1	Wed, 03 Jun 2015 15:57:33
2961	↓357	Robinsheridan99	0.50718	1	Thu, 18 Jun 2015 15:34:04
2962	↓357	Dill_Pickle	0.50239	1	Fri, 15 May 2015 04:18:51
2963	↓357	NicolasBurl	0.50239	4	Sun, 14 Jun 2015 13:37:51 (-0.7h)
2964	↓357	gouthambilakanti	0.49761	1	Tue, 16 Jun 2015 06:50:44
2965	↓357	Team-Robert	0.49282	1	Thu, 14 May 2015 08:34:49
2966	↓357	Zeger Hoogeboom	0.47368	1	Sun, 17 May 2015 19:00:12
2967	↓357	Alex Miller	0.46890	1	Wed, 29 Apr 2015 05:17:52
2968	↓357	GrishaSizov	0.40670	1	Sun, 24 May 2015 15:19:46
2969	new	Brendy Doherty	0.37799	2	Sun, 21 Jun 2015 20:51:11 (-0.4h)
2970	↓358	Easylove	0.37321	1	Thu, 14 May 2015 08:49:40
2971	new	Chaoran Wei	0.23445	1	Wed, 24 Jun 2015 19:38:44
2972	new	Carol	0.00000	1	Sun, 21 Jun 2015 19:55:29

Download raw data

Figure 3.3: Snapshot of kaggle competition "Titanic: Machine Learning from Disaster" leaderboard last screen

Classification tree was made on the dataset of users' smartphone to predict the user activity, which resulted in 82% correctness. Random forest analysis on the same dataset improved the prediction precision by 8% and became 90%. Moreover, participation in kaggle competition for the same ranked the model at 21st among 30 participants [12], as shown in Figure 3.4 and Figure 3.5.

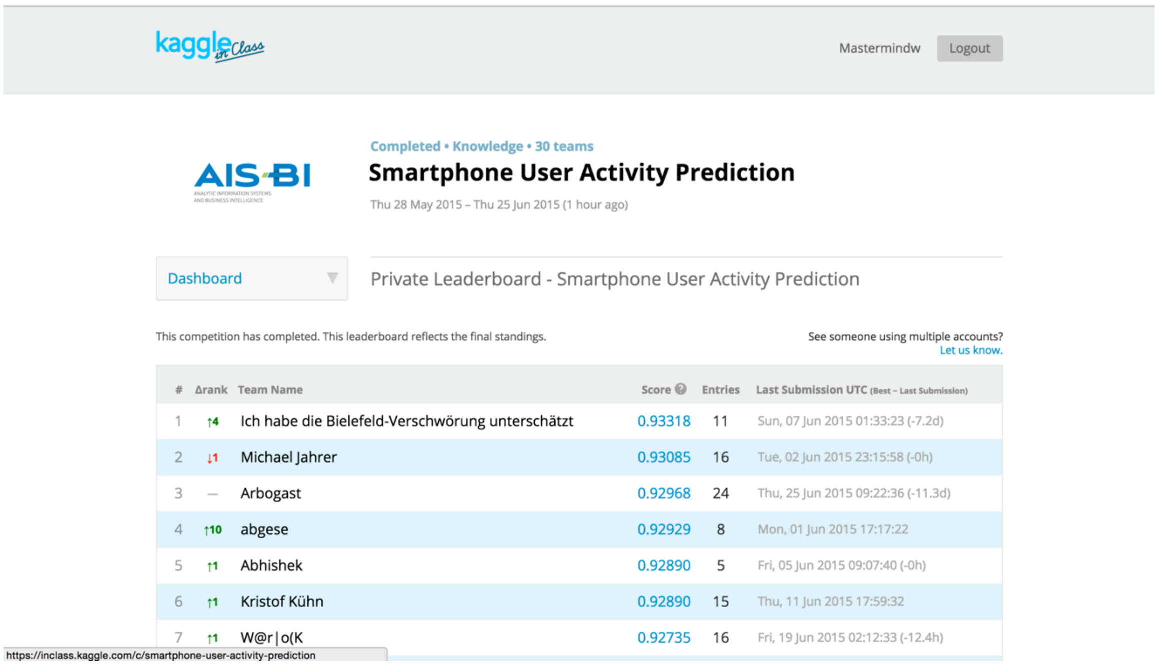


Figure 3.4: Snapshot of kaggle competition "Smartphone User Activity Prediction" leaderboard initial screen

18	↑1	xiaoge	0.91686	6	Thu, 25 Jun 2015 12:51:01 (-3h)
19	↓1	piotrek	0.91064	13	Sat, 20 Jun 2015 22:06:36 (-9d)
20	—	DL	0.90754	6	Sun, 14 Jun 2015 20:39:54 (-8.3h)
21	—	**Mastermindw**	**0.90754**	**8**	**Thu, 25 Jun 2015 15:53:04 (-5h)**
Your Best Entry ↑ Your submission scored **0.90559**, which is not an improvement of your best score. Keep trying!					
22	—	nonagon	0.90676	7	Fri, 05 Jun 2015 01:25:12
23	—	Tomas Vantuch	0.89744	7	Fri, 12 Jun 2015 06:02:07 (-7.8d)
24	—	Timur Sattarov	0.87568	3	Thu, 04 Jun 2015 22:23:22 (-6.2d)
25	—	Daneil	0.74709	3	Thu, 11 Jun 2015 19:00:59
26	—	Priyansu	0.19425	11	Thu, 25 Jun 2015 17:59:07 (-2.3h)
		All 1s Benchmark	0.15890		
27	—	inversion	0.15890	1	Thu, 28 May 2015 15:08:44
28	—	LouisZheng	0.15890	1	Tue, 23 Jun 2015 09:50:38 (-1.8h)
29	—	ilarele	0.00000	2	Tue, 02 Jun 2015 19:23:44 (-0h)
30	—	SanchitChiplunkar	0.00000	2	Mon, 15 Jun 2015 05:05:29 (-0.1h)

Download raw data

Figure 3.5: Snapshot of kaggle competition "Smartphone User Activity Prediction" leaderboard rank and last screen

Random forest technique was found to be the most efficient to make prediction on the selected datasets.

CHAPTER 4

4. CONCLUSION AND FUTURE WORK

The proposed model still leaves much room for improvement. More factors can be analyzed to improve the produced model to increase the prediction precision. For instance, survival rate of passengers can be improved by taking into account the cabin numbers of the passengers because it describes data analyst how close or far the individual is from the split area and exits. Similarly, user activity from smartphone prediction can also be improved by applying different statistical modeling techniques with increased number of iterations and large data set.

Moreover, with the rise of sun every single day there are more and more datasets that are being analyzed by data analysts around the world. Studied datasets were just to get the hands dirty, since there is a plethora of datasets to conclude and get the answers to real life problems.

5. REFERENCES

1. V. Dhar, 'Data Science and Prediction', *Cacm.acm.org*, 2013. [Online]. Available: http://cacm.acm.org/magazines/2013/12/169933-data-science-and-prediction/fulltext. [Accessed: 24- Apr- 2015].
2. J. Leek, 'The key word in “Data Science” is not Data, it is Science | Simply Statistics', *Simplystatistics.org*, 2013. [Online]. Available: http://simplystatistics.org/2013/12/12/the-key-word-in-data-science-is-not-data-it-is-science/. [Accessed: 24- Apr- 2015].
3. Villanova University, 'Careers in Big Data'. [Online]. Available: http://www.villanovau.com/university-online-programs/careers-in-big-data/. [Accessed: 24- Apr- 2015].
4. AOL Jobs, 'Data Scientist: The Hottest Job You Haven't Heard Of', 2011. [Online]. Available: http://jobs.aol.com/articles/2011/08/10/data-scientist-the-hottest-job-you-havent-heard-of/. [Accessed: 24- Apr- 2015].
5. 2012. [Online]. Available: http://www.insideactivitytracking.com/data-science-activity-tracking-and-the-battle-for-the-worlds-top-sports-brand/. [Accessed: 24- Apr- 2015].
6. Archive.ics.uci.edu, 'UCI Machine Learning Repository: Human Activity Recognition Using Smartphones Data Set', 2012. [Online]. Available: http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones. [Accessed: 24- Apr- 2015].
7. Reyes-Ortiz, Jorge Luis et al. 'Human Activity And Motion Disorder Recognition: Towards Smarter Interactive Cognitive Environments.'. 21st European Symposium On Artificial Neural Networks, Computational Intelligence And Machine Learning, ESANN 2013. Available: https://www.elen.ucl.ac.be/Proceedings/esann/esannpdf/es2013-11.pdf. [Accessed: 10- Apr- 2015]
8. Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier. Parra, and Jorge-Luis. Reyes-Ortiz. ‘A public domain dataset for human activity recognition using smartphones’. European Symposium on Artificial Neural Networks, Computational Intelligence and

Machine Learning (ESANN), 2013.
Available:https://www.elen.ucl.ac.be/Proceedings/esann/esannpdf/es2013-84.pdf. [Accessed: 10- Apr- 2015]

9. Anguita, Davide et al. 'Human Activity Recognition On Smartphones Using A Multiclass Hardware-Friendly Support Vector Machine'. 4th *International Workshop Of Ambient Assisted Living (IWAAL 2012)*. 2012. Available: http://www.icephd.org/sites/default/files/IWAAL2012.pdf[Accessed: 17- Apr- 2015]
10. Anguita, Davide et al. 'Energy Efficient Smartphone-Based Activity Recognition Using Fixed-Point Arithmetic.'. *Journal of Universal Computer Science* 19.9 (2013) Available:http://www.jucs.org/jucs_19_9/energy_efficient_smartphone_based/jucs_19_09_1295_1314_anguita.pdf. [Accessed: 17- Apr- 2015]
11. Kaggle Titanic. Available:https://www.kaggle.com/c/titanic [Accessed: 20-May-2015]
12. Kaggle Smartphone. Available:https://inclass.kaggle.com/c/smartphone-user-activity-prediction [Accessed: 02-Jun-2015]

6. PLAGARISM REPORT (TURNITIN DIGITAL RECEIPT)

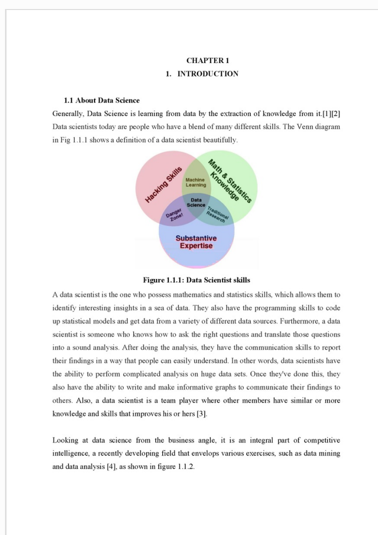
turnitin

Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Waqas Ahmed
Assignment title:	Final Report
Submission title:	Final Resubmission
File name:	IS_for_turnitin_2.docx
File size:	269K
Page count:	11
Word count:	1,633
Character count:	8,658
Submission date:	21-Jul-2015 12:49PM
Submission ID:	555115025



CHAPTER 1

1. INTRODUCTION

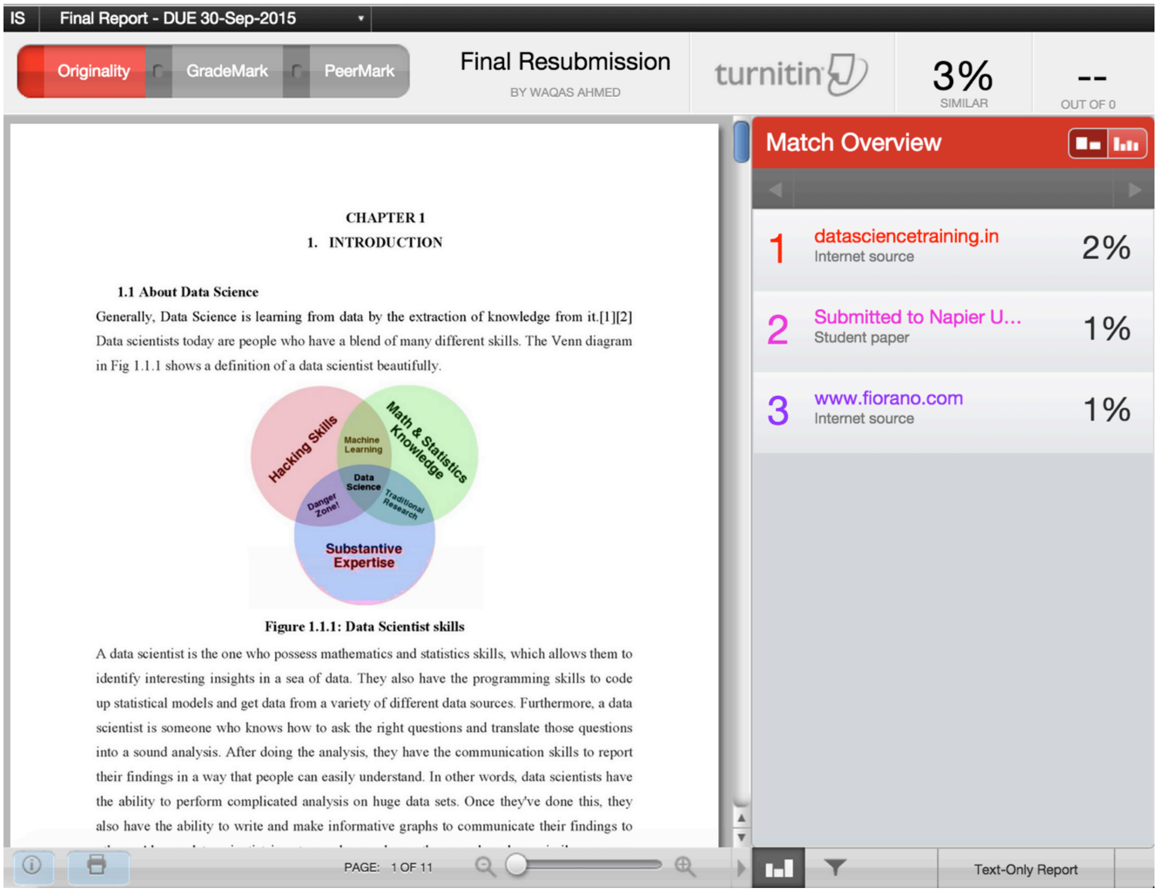
1.1 About Data Science

Generally, Data Science is learning from data by the extraction of knowledge from it.[1][2] Data scientists today are people who have a blend of many different skills. The Venn diagram in Fig 1.1.1 shows a definition of a data scientist beautifully.

Figure 1.1.1: Data Scientist skills

A data scientist is the one who possess mathematics and statistics skills, which allows them to identify interesting insights in a sea of data. They also have the programming skills to code up statistical models and get data from a variety of different data sources. Furthermore, a data scientist is someone who knows how to ask the right questions and translate those questions into a sound analysis. After doing the analysis, they have the communication skills to report their findings in a way that people can easily understand. In other words, data scientists have the ability to perform complicated analysis on huge data sets. Once they've done this, they also have the ability to write and make informative graphs to communicate their findings to others. Also, a data scientist is a team player where other members have similar or more knowledge and skills that improves his or hers [3].

Looking at data science from the business angle, it is an integral part of competitive intelligence, a recently developing field that envelops various exercises, such as data mining and data analysis [4], as shown in figure 1.1.2.


IS
Final Report - DUE 30-Sep-2015
Originality
GradeMark
PeerMark
Final Resubmission
BY WAQAS AHMED
turnitin
3%
SIMILAR
--
OUT OF 0
CHAPTER 1
1. INTRODUCTION
1.1 About Data Science
Generally, Data Science is learning from data by the extraction of knowledge from it.[1][2] Data scientists today are people who have a blend of many different skills. The Venn diagram in Fig 1.1.1 shows a definition of a data scientist beautifully.
Hacking Skills
Math & Statistics Knowledge
Machine Learning
Data Science
Danger Zone!
Traditional Research
Substantive Expertise
Figure 1.1.1: Data Scientist skills
A data scientist is the one who possess mathematics and statistics skills, which allows them to identify interesting insights in a sea of data. They also have the programming skills to code up statistical models and get data from a variety of different data sources. Furthermore, a data scientist is someone who knows how to ask the right questions and translate those questions into a sound analysis. After doing the analysis, they have the communication skills to report their findings in a way that people can easily understand. In other words, data scientists have the ability to perform complicated analysis on huge data sets. Once they've done this, they also have the ability to write and make informative graphs to communicate their findings to
PAGE: 1 OF 11
Match Overview
1 datasciencetraining.in Internet source 2%
2 Submitted to Napier U... Student paper 1%
3 www.fiorano.com Internet source 1%
Text-Only Report

Turnitin Originality Report

Final Resubmission by Waqas Ahmed

From Final Report (IS)

turnitin

- Processed on 30-Jul-2015 00:33 PKT
- ID: 555115025
- Word Count: 1633

Similarity Index
3%
Similarity by Source

Internet Sources:
3%
Publications:
0%
Student Papers:
1%

sources:

1 2% match (Internet from 22-Feb-2015)
http://datasciencetraining.in/data-science-courses-john-hopkins-university-free-course-coursera/

2 1% match (student papers from 05-Apr-2012)
Submitted to Napier University on 2012-04-05

3 1% match (Internet from 21-May-2014)
http://www.fiorano.com/docs/fsoa/921/pdf/fioranoestudiouserguide.pdf

paper text:

CHAPTER 1 1. INTRODUCTION 1.1 About Data Science Generally, Data Science is learning from data by

7. RESUME

WAQAS AHMED

FULL STACK SOFTWARE ENGINEER

WORK EXPERIENCE: 6 YEARS

Microsoft CERTIFIED Professional

ORACLE Certified Expert Oracle Database SQL

Microsoft CERTIFIED Technology Specialist .NET Framework 4, Web Applications

Microsoft Specialist Programming in HTML5 with JavaScript and CSS3 Specialist

Cell: +971-56-2418355

E-mail: wakqasahmed@gmail.com

OBJECTIVE

TO BE PART OF THE TEAM WHO COULD MAKE SIGNIFICANT DIFFERENCE USING LATEST TECHNOLOGIES.

WORKING EXPERIENCE

- **Full Stack Software Engineer** — **Aug 2012 – Present**
  Trix Communications (Dubai, United Arab Emirates)
- **Software Engineer** — **Jan 2011 – Jul 2012**
  TPS Pakistan Pvt. Ltd. (Karachi, Pakistan)
- **Junior Software Developer** — **Jun 2010 – Jul 2011**
  C.R.E.A.M. Solutions (Karachi, Pakistan)
- **Chief Design Officer** — **Aug 2009 – Jan 2011**
  The Maverickz FZC (Co-Founder) (Karachi, Pakistan)
- **Junior Software Engineer** — **Jun 2009 – Aug 2009**
  2 Steps Solutions (Karachi, Pakistan)

TECHNICAL SKILLS

- **Languages:**
  C#, ASP.NET, PHP, Actionscript, MXML, SQL, HTML5, CSS3, Javascript, JQuery, AJAX, JSON
- **Database: Oracle, MySQL**
- **Platforms: Microsoft Windows**, Red Hat Linux
- **Application Frameworks: ASP.NET** & **MVC, Cairngorm (Flex), CodeIgniter, Twitter Bootstrap**
- **Content Management Systems: Wordpress, Sitefinity**
- **eCommerce: Magento**
- **Version Control: Git, Visual Sourcesafe**
- **Others: Crystal Reports**, VMWare, Search Engine Optimization, **Windows Services, Web Services, Responsive Web Design, Web Optimization**, Gamification, **Facebook Applications**

Recent: MEAN (mongo, express, angular, node.js) Stack, Phonegap

ACADEMIC QUALIFICATIONS

Dubai, U.A.E. Sep 2014 – Sep 2016	SZABIST **M.S., Software Engineering**
Karachi, Pakistan Jan 2007 – Dec 2010	U.B.I.T., University of Karachi **B.S., Computer Science**
Karachi, Pakistan Jan 2005 – Dec 2006	Govt. College Formen Nazimabad **Intermediate** Faculty: Science Pre-Engineering
Karachi, Pakistan Jan 2002 – Dec 2003	Metropolitan School **Matriculation** Faculty: Computer Science

PERSONAL INFORMATION

Father's Name: Abdul Rasheed
Visa Status: Employment Visa (Dubai, UAE)
Nationality: Pakistan
Marital Status: Married

OTHER SKILLS

Good Communication Skills	Responsible
Energetic and Dynamic team player	Creative and Imaginative
Proficient in English Language (Verbal & Non-Verbal)	