Pradeep Thalasta

2652 Ellendale Place, Apartment 1,
Los Angeles, California, 90007.

+1(323) 690-6523 | thalasta@usc.edu | prodo56 | pradeepthalasta5693

Education

University of Southern California, Los Angeles, California
GPA: 3.6

M.S. in Computer Science
August. 2016 - May. 2018

Analysis of Algorithms, Foundations of Artificial Intelligence, Applied Natural Language Processing, Database System.

National Institute of Technology, Karnataka, Surathkal, India
CGPA: 8.63

B.Tech. in Information Technology
July. 2011 - May. 2015

Data Structures and Algorithms, Data Warehousing And Data Mining, Advanced Database Systems, Web Technologies, Web Services, Semantic Web

Skills

Languages

Java, C, C++, Python, SQL, Shell scripting, JavaScript

Applications/Framework

Hadoop, Map Reduce, Tensorflow, Apache Spark, Numpy, Scikit-learn, Pandas, Weka, R, Cassandra, Redis, Elasticsearch, Kibana, Logstash, RabbitMQ, Apache Lucene, Apache Kafka, Spring, Jenkins, Maven, Hibernate, Eclipse, Git, MySQL, Oracle 11g

Platform and Operating Systems

Amazon Web Services, Google Cloud Platform, Microsoft Azure, Unix, Linux, Mac OSX, Windows

Experience

Coefficient Labs
Los Angeles, USA

Data Science Intern
June. 2017 - Present

- Regression Analysis of Data for New Clients
- Implemented a Recommendation system of new Ads using Predictive Analysis based on Linear Regression

Samsung R&D Institute India-Bangalore
Bangalore, India

Software Engineer
July. 2015 - July. 2016

- Worked on Server side programming and designed the service architecture as a part of Internet of Things and Smart Assistant Team for ***Jifical*** application(available on Play Store).
- Implemented applications based on **Machine Learning, Natural Language Processing**, MVC design pattern for the server on AWS using different frameworks like **Spring MVC, Hibernate, Java, Cassandra, Redis, Apache Kafka, and ELK**.
- Implemented a **Log Processing Engine** for analytical purpose for the designed application based on Logstash and Elasticsearch
- Developed a Dashboard based on Kibana to query the data for visualisation and analytics based on Log information stored in Elasticsearch.
- Filed ***Patent*** Application on Automatically managing operations of electronic device based on user input and grip pattern.
- Obtained ***Most Promising Idea Award*** during intra division Hackathon.

Projects

Speech Based Summarisation and Emotion detection
Los Angeles, USA

Natural Language Processing
March. 2017-April. 2017

- Implemented Speech Based Summarisation and Emotion detection in python for German Language.
- Built a training model based on ***Naive Bayes, Support Vector Machines, Logistic Regression and Neural Networks*** to detect various emotions like Anger, Sadness, Anxiety, Boredom, Neutral, Disgust, and Happiness.
- Used various features like Energy, Fundamental Frequency, Loudness, and Chromograms to train the model and to classify the given audio sample for various emotions.
- Implemented ***TextRank*** algorithm to Summarise the content of the audio input.

Quora Answer Prediction using Apache Spark
Los Angeles, USA

Machine Learning and Natural Language Processing
June. 2017 - June. 2017

- Implementing a code to predict if a given question will be answered within a day given the topic and the number of followers.
- Implementation requires cleaning and preprocessing of the dataset followed by Natural Language Processing techniques like ***word2vec*** and techniques to find the similarity between the question sets.
- Implemented code based on ***Logistic Regression*** to classify the given question
- Implementation requires knowledge about python, Apache-Spark, Machine Learning techniques to classify each question accordingly.

Dialogue System

Los Angeles, USA

Directed Research in Natural Language Processing, Machine Learning and Information Retrieval — *Feb. 2017 - April. 2017*

- Implemented a simple ***Information Retrieval based Question Answering System*** as a part of Directed Research under the guidance of Prof. Kevin Knight (Research Director, Natural Language Technologies and Fellow, Information Sciences Institute).
- Data was collected from various sources like blogs, twitter, interview data, etc. using web crawlers and scrappers.
- Implementation involved cleaning of data collected, storing of data in Databases for later retrieval using ***Jaccard similarity coefficient***.
- Used tools based on Apache Lucene, python, and NLTK for implementation.

Parts of Speech(PoS) Tagging using Hidden Markov Model

Los Angeles, USA

Natural Language Processing — *March. 2017*

- Implemented ***Hidden Markov Model*** to tag Catalan Language corpus.
- Corpus was adapted from the Catalan portion of WikiCorpus v. 1.0
- Implemented the code in python using ***Viterbi Decoding Algorithm*** to tag the test data
- Obtained an accuracy of 90%

Sentiment Analysis of Hotel Reviews

Los Angeles, USA

Natural Language Processing — *January. 2017*

- Implemented ***Naive Bayes Model*** for analysing the reviews in python.
- Implemented word tokeniser using regular expressions to extract meaningful words from the corpus.
- Reviews were classified as truthful positive, fake positive, truthful negative and fake negative.
- Obtained F1 score of 0.94.

Web Service Discovery and Recommendation

Surathkal, India

Researcher for Web Service Discovery. — *July. 2014 - May. 2015*

- Designed and implemented Web Service Discovery and Recommendation to the user.
- Involved skills on ***Machine Learning and Natural Language Processing*** using Python and MongoDB.
- Enhanced the efficiency and correctness based on clustering and classification of web services.

A Graph Based Technique to Find Common News Contents From Multiple News Sources

Surathkal, India

Researcher for Data Mining and Information Retrieval — *January. 2014 - April. 2014*

- Designed and implemented a graph based algorithm to find the common news contents.
- Involved skills on ***Natural Language Processing, Machine Learning and Data Mining*** using Java, MySql and web crawlers.
- Improved the correctness of the results using graph based techniques.
- Enhanced and proposed a new method to find common news contents from different sources.

Honors & Awards

2016	**Patent**, Samsung R&D Institute India-Bangalore.	*Bangalore, India*
2016	**Most Promising Idea Award**, Samsung R&D Institute India-Bangalore intra-division Hackathon	*Bangalore, India*
2016	**1st Place**, Samsung R&D Institute India-Bangalore intra-division Hackathon	*Bangalore, India*
2016	**1st Place**, Facebook Messenger Bot Challenge	*Los Angeles, USA*
2014	**6th Place**, Microsoft-Hackon Ethical Hacking Competition (National Level).	*Hyderabad, India*

Program Committees

2017-2018	**Vice President**, ASSOCIATION FOR THE ADVANCEMENT OF ARTIFICIAL INTELLIGENCE at USC	*Los Angeles*

Online Courses

2017	**Google**, Deep Learning	*Udacity*
2016	**Stranford University**, Machine Learning	*Coursera*
2017	**University of California, San Diego**, Hadoop Platform and Application Framework	*Coursera*
2016	**University of Illinois at Urbana-Champaign**, Text Mining and Analytics	*Coursera*
2017	**Microsoft - DAT210x**, Programming with Python for Data Science	*edx*
2017	**Microsoft: DAT202.3x**, Implementing Predictive Analytics with Spark in Azure HDInsight	*edx*