

UMANG MEHTA

Location: Houston, TX
umangrmehta.github.io

Email: umangrmehta@gmail.com
linkedin.com/in/umangrmehta

Cell: +1 (812) 955-9507
github.com/umangrmehta

EDUCATION

Master's Degree in Data Science

MAY 2019

School of Informatics Computing and Engineering (SICE), Indiana University, Bloomington

Major Courses: Advanced NLP, Machine Learning, Data Mining, Social Media Mining, Elements of AI, Deep Learning Systems. GPA: 3.6

Bachelor's Degree in Computer Engineering

MAY 2013

K J Somaiya College of Engineering (KJSCE), Mumbai University, India (GPA: Not on a 4.0 Scale)

WORK EXPERIENCE

Hello Nesh Inc.(Nesh), Houston, TX

JUNE 2019 - OCT 2019

Data Scientist - NLP & AI

- Built Conversational AI backend using Rasa framework in Python for Oil and Gas domain
- Developed Knowledge Extraction pipeline for public documents with Semantic Extraction using NER, Constituent Parsing with Rasa NLU & Spacy and built a Knowledge Graph using Dgraph, GraphQL to represent the extracted knowledge
- Extracted topics from public documents using Gensim & TextRazor and linked them to entities in Knowledge Graph
- Conceptualized and built PoC of Diagnostic and Predictive Analysis Engine for Oil Well Failure using E-M, MLE and MAP
- Trained and deployed text classifiers with Word Embeddings, LSTM, BERT using TensorFlow, Keras on AWS Sagemaker

Kelley School of Business at Indiana University, Bloomington, IN

JUNE 2018 - MAY 2019

Graduate Research Assistant under Prof. Matthew Josefy

- Research and implement NLP methods to extract relevant information from SEC Filings using Python
- Develop or implement models using ML and NLP to analyze the business model and board leadership structure of companies
- Built Text Classifiers using SciKit-Learn with around 90% accuracy measured with 10-Fold Cross Validation

AriaData Inc. (Aridat), Bloomington, IN

FEB 2018 - MAY 2019

Chief NLP Research Engineer (Unpaid Position)

- Build an analytics engine to determine the critical reception of an artists work based on chatter on social media using Python
- Leading the research and implementation of advanced NLP methods to improvise the analytics
- Developed a Data Preprocessing module using NLTK with steps involving POS Tagging, Stop Words Removal, Stemming and Lemmatization, Negation Handling, N-gram and Sentiment Scoring using AFINN and TextBlob for feature engineering
- Built Sentiment Classifier using an ensemble of Naïve Bayes and Logistic Regression to classify tweets with 87.89% accuracy

Vitruvian Technologies Pvt. Ltd. (Realty Redefined), Mumbai, India

JUNE 2014 - JULY 2017

Senior Developer & Team Mentor

- Collaborated in a team of 12 for project development including Java and Scala Programming, Data Structure & Database Design in MySQL, Web Design & Development and Unit Testing for web-based ERP & CRM systems in Real Estate domain
- Lead, trained and mentored a sub-team of 5 throughout the development of the projects

Algonation, Mumbai, India

JUNE 2013 - MAY 2014

Co-founder, Developer

- Developed websites, web portals and mobile apps for small and medium scale enterprises using Java and PHP
- Built server software for customized TCP Layer Protocols for cloud-based industrial requirements using Java
- Mentored and trained multiple groups of 3-4 undergraduate interns for developing industry level projects

Research Innovation Incubation Design Laboratory (Riidl), Mumbai, India

JUNE 2012 - MAY 2013

Software Engineer Intern

- Developed web-based ERP application for educational institutes using HTML, CSS, Bootstrap, Javascript and PHP
- Designed the data structures & schema and managed the database transactions using MySQL
- Built mobile app for the ERP using Java Android SDK for Android phones and Java ME for Java-based feature phones

TECHNICAL SKILLS & CERTIFICATIONS

Languages	: Java (5 yrs), Javascript (5 yrs), Scala (1 yr), Groovy (1 yr), Python (1 yr), C/C++, R, MATLAB, GraphQL, Gremlin
Database	: RDBMS- MySQL(5 yrs), PostgreSQL, Oracle; NoSQL- MongoDB, Neo4J, JanusGraph, Dgraph, Apache Solr(1 yr)
Frameworks	: TensorFlow, Keras, DL4J, ND4J, Scikit-Learn, NumPy, SciPy, Pandas, Flask, Dash(Plotly), Spring, Hibernate
NLP & Big Data	: NLP- Rasa, NLTK, Stanford CoreNLP, Spacy, OpenNLP, Spark NLP, Polyglot, Gensim; Big Data- Hadoop, Spark
Certification	: Oracle Certified JavaSE6 Programmer (OCPJP)
Servers	: WildFly 15, Apache 2, Apache Tomcat 7 & 8, NodeJS, Netty, Openfire, Nginx
Deployment	: Docker, Kubernetes, AWS EC2, AWS Lambda, AWS Sagemaker, Google Cloud Compute Engine
OS & Cloud	: Ubuntu(Linux), Windows, macOS, Amazon Web Services(AWS), Google Cloud Platform

LEADERSHIP & EXTRA-CURRICULAR

- Elected and worked as the President and Treasurer of the Data Science Club at Indiana University
- Organized hackathon for analyzing the opioid crisis in collaboration with SPEA at Indiana University
- Won 2 consecutive Annual Fall Chili Cook-offs in 2017 and 2018 organized by SICE at Indiana University
- Organized Technical and Cultural Festivals in KJSCE as Creative Head of the Organizing Committee