

Siddhant Gawsane

Data Scientist
F1-OPT EAD

Dallas/New York

☎ (682) 203-8916

✉ siddhant.gawsane@mavs.uta.edu

g github.com/siddhantgawsane

in [linkedin.com/in/siddhantgawsane](https://www.linkedin.com/in/siddhantgawsane)

RESEARCH EXPERIENCE

IDIR Labs at UTA, Arlington Tx — *Research Assistant*

DEC 2015 - SEPT 2016

ClaimBuster – SVM based algorithm classifies a sentence into ‘fact’ or ‘non-fact’ using NLP, text mining, sentiment analysis and probabilistic modelling

Live stream demos for the 2016 Presidential debates using closed captions

Organizing, pulling, cleaning, munging and storing training data. Data Visualizations in D3.js, chart.js, tableau

EDUCATION - MASTERS IN COMPUTER SCIENCE

The University of Texas, Arlington Tx — *Master Of Science*

DEC 2017 (anticipated), FOCUS - Data Science/Artificial Intelligence, GPA 3.65

Machine Learning

‘News Groups’ for news articles using tf-idf, naive bayes – 74% accuracy

Species of the iris plant using linear regression, KNN cluster– 88.8% accuracy

Dimensionality reduction experiments on sparse datasets using PCA, LDA like approach

Neural Networks

Building auto-encoders for the pen digits problem using Keras, Theano and from scratch, and testing ease of use, RMSE and running time on each

Artificial Intelligence

Comparing Bayesian Networks, Decision tree/forest, Naive Bayes, KNN approaches to predicting Pen Digits, LandSat, Yeast datasets - 57-84% acc
Game play, Minimax, a/b pruning

Data Mining

Tf-Idf text mining based document search based on a corpus of election debates from 1960-2012

Kaggle competition Home Depot product suggestions based on search query with score 0.57 and rank 129

Hadoop map-reduce queries for department wise sales data, web-logs etc

Adv. Topics in Scalable Image Search

SIFT/SURF/ORB based feature extraction and object recognition projects

Web And Social Analytics

Text Summarization – Extracting actionable insights from Yelp reviews using k-means clustering. Pointing out top 10 most frequent topics for a restaurant. ROI upto 750\$/month

SKILLS

Deep learning, machine learning, data mining, data visualizations

Map reduce, bayesian statistics, model training, data fitting, test/train split, feature engineering, data gathering

Bayesian statistics, hypothesis tests, statistical tests, distributions, maximum likelihood estimators

Applied mathematics, statistics, probability theory, related fields

Kernel learning, SVM, bayesian networks, probabilistic modeling, neural networks, CNN, CPU/GPU, p-value estimates, Kmean clustering, gaussian processes, k-nearest neighbors, random forests, ensemble methods

Theano, Keras, OpenCV, NLTK, scikit-learn, LibSVM, Pandas, Matplotlib, D3.js, NumPy, SciPy

Hadoop, Spark, Flink, PIG

AWS, GCP, MS Azure, IBM Bluemix

ES6, NPM, Bootstrap, JQuery, Knockout, React, Backbone, Angular, Cocoa, Ember, Node

Confluence, JIRA, Jenkins, Slack, Maven, Bugzilla, Cassandra, JUnit

TDD, Jasmine, Cucumber, JSLint

Atom, Sublime, IntelliJ, Vim

Git, SVN, perforce, linux shell

PUBLICATIONS

AAAI 2017 - ClaimBuster: Monitoring Political Discourses and Spotting Factual Claims

VLDB 2017 - ClaimBuster: The First-ever End-to-end Fact-checking System

KEY COURSES

Data Analysis and Modelling

Techniques - Statistical tests, distributions, maximum likelihood estimators, using Linear Algebra and Multi-Variable Calculus, data-driven approach in predictive models

Cloud Computing - Developing cloud-based web apps on AWS EC2, IBM Bluemix, Microsoft Azure and Google Cloud Platform

Pune University, Pune, India — *Bachelor of Engineering*

MAY 2013, FOCUS - Information Technology, GPA 3.41

WORK EXPERIENCE - 5+ Years

Mouser Electronics, Mansfield Tx — *Info Services Intern*

JUNE 2016 - DEC 2016

Collaborating with stakeholders, UX Designers, testers on a close knit team

Advocating UX best practices, migrating business logic to front-end.

Building SPAs, stand-alone widgets in KnockoutJS, requireJS, underscoreJS

AJAX/JSON REST API on Django/Python, NOSQL database objects MongoDB

eQ Technologic, India — Software Developer

APRIL 2014 - JULY 2015

Building a data warehousing and business intelligence reporting and snapshots tool that conceptualizes data into cubes and performs filters, slices, drill-up/down, rotations and skews

Mobile responsive, scalable, stand-alone, test-driven widgets in BackboneJS, UnderscoreJS, JQuery, RequireJS, Jasmine, KendoUI, BootstrapJS

Building a widget framework library for cross-product use which acts as widget factory, to centralize initialization of backboneJS widget objects

Code quality, page rendering, lazy loading, mobile optimization, caching and uglification using JSLint, IntelliJ, JSMIn, uglify2, GZIP compression

Adaptavant Tech, India — Jr. Software Engineer

MAY 2013 - APRIL 2014

J2EE based web application using JQuery, Javascript, Spring 3.0, SOAP/XML, REST/JSON, AJAX, JDBC, MySQL, Oracle, EJB, POJO, Maven, ANT, Tomcat etc.

Training on NPM libraries and MV* front-end frameworks AngularJS, UI elements on BootstrapJS, HTML5, CSS3

CarlQ, India — Intern

NOV 2012 - MAY 2013

Native Android programming, Eclipse IDE and Android SDK, Google play services and Google Maps API V2 integration

JSP servlet, RPC based port programming, wireless bluetooth communication

Web Data Management -

Developing Ruby on Rails and NodeJS web-apps on NOSQL data

Secure Programming -

Analyzing the most common security flaws in web pages and their fixes

Adv Software Engineering -

Building web applications that can withstand performance tests in the face of big data

WEB ARCHITECTURES

HTTP/REST

Single Page Apps

Widget Framework

Responsive Layouts/Bootstrap

3-Tier Architecture

Object Relational Mapping

AJAX/JSON

SOAP/XML

PROCESSES

Continuous Integration/Delivery

Agile Model

Pair Programming

Lean Development

LANGUAGES

Javascript, Java, Python, Ruby, C/C++

INDEPENDENT PROJECTS

UTA Rideshare - ride sharing app
uta-rideshare.appspot.com