Kavita Ganeshan

Personal Data

Address:	B-701, Pranay Sudarshan Apts, Joshi lane Ghatkopar East, Mumbai - 400077, India
Phone:	+91-9920552210
nationality	Indian
Date of birth	08.04.1987
email:	kavlata@gmail.com
github:	https://github.com/kavlata
linkedin:	http://www.linkedin.com/pub/kavita-ganeshan/22/616/711

Education

June 2010	MTech in Information Technology, **IIIT**, Bengaluru CGPA: **3.33/4.00**
June 2008	Btech in Computer Science, **PVPPCOE**, Mumbai GPA: **72**%
2003-2004	HSC **Maharashtra Board**, Mumbai GPA: **82.50**%
2001-2002	SSC **Maharashtra Board**, Mumbai GPA: **89.20**%

Summary

I am Senior Data Scientist at Cuddle.ai-Fractal Analytics Pvt Ltd. I have 7 years of industrial experience in Natural Language Processing and Machine Learning.

My current role is in the cross-roads of recommendation , ranking algorithms, software engineering, machine learning, data science and natural language processing

Work Experience

Current Sep 2015	Senior Data Scientist-NLP at Fractal Analytics Pvt Ltd - Cuddle.ai Mumbai, India *News Recommendation* I have designed and developed an engine to recommend News to users based on their business interests. I have setup a pipeline using word2vec based relevance module, minhash-LSH based deduplication module, collborative filtering based recommendation and ranking based on the feedback stats from Redshift AWS. The News digest is being used by 50 users within the company and 200 client users who have been actively participating for over 7 months. My ensemble algorithm has increased the accuracy of Relevance by 60 percent from our baseline. My recommendation engine has led to a 3 percent increase in user engagement calculated based on user's interactions. I am also involved in building automated contextual ontology for the business and to enrich it with insights extracted from News
Apr 2015 - Sep 2015	Senior Software Engineer at Germinait Solutions Pvt Ltd, Mumbai, India *Algorithms Team* I was part of Algorithms team and my role was to extract user intent from tweets. I had implemented one-vs-all multiclass SVM classifier. Daily the engine would process 10 batches of data, each of size 20 MB. The sparsity of the data was - 1 user intent per 600 tweets (approx). My algorithm performed with an accuracy of 0.8

Aug 2011 - Mar 2015	Senior Software Engineer at Creditpointe Services Pvt Ltd, Pune, India
	Semantic Team
	I was part of Semantics Team. I had worked on query interpretation, semi-supervised context based sentiment analysis (using NLP techniques), topic modelling using the grammar in the corpora,semi-supervised ontology generation, automatic discovery of ontology using Wikipedia, relation extraction, sentiment analysis, logical deduplication and co-reference resolution. I deployed above solutions for Financial, Pharma and Legal contracts domain.

Aug 2010 - Aug 2011	Information Retrieval Engineer at Web 18 Pvt ltd Mumbai, India
	I was part of Information retrieval team and worked on query segmentation, extracting quotations mentioned in news articles, Named entity extraction, classifying News articles and clustering similar content.

Dec 2009 - Jun 2010	Internship at GISE lab , IITB
	I worked under Prof.N.L.Sarda in the GISE lab in IIT Bombay, India. My work helped me publish couple of papers in reputed journals.
	I had presented my internship work in TCS innovation labs under the supervision of Dr.Arun Pande. (Head of TCS innovation labs, Mumbai, India)

Data Challenge

CODS 2017	Differentiating Code-Borrowing from Code-Mixing.
	I was a runner up in this data challenge. I submitted this as a team entry. I presented this work in CODS/COMAD 2017 conference in IITM. Details are here https://cuddledatascience.github.io/CODS2017/

Scholarships and Publications

Nov 2010	Keyword Search in geo-spatial databases
	Submitted in demo paper category in ACM SIG SPATIAL 2010, USA
Sep 2010	Developing IITB Smart CampusGIS Grid
	Submitted in paper category in A2CWIC 2010, India
2005-2006	Awarded the Sir Ratan Tata Scholarship for academic excellence in 2nd Year bachelors program
2012-2013	Received Best Employee award in Creditpointe Services.

Computer Skills

Languages:	Java, Python
NLP/ML Libraries:	Stanford Parser, NER, Coreference resolution, Wordnet, sentiwordnet, conceptnet, semgrex , Opennlp, Wikipedia
Frameworks/Libraries/Tools:	Lucene, Solr, LibSVM, Word2Vec
Databases :	MySQL, PostgreSQL
Editors:	vi
Cloud:	AWS - SQS, EC2, ECS. Microsoft cognitive API - text
Continuous Integration:	Jenkins, GIT, Docker
Others:	Latex, Intellij
Operating Systems:	Mac, Linux
Experimented with:	R,Python - Scikitlearn, nltk, Gensim,Numpy, Pandas Neo4j, MongoDB
Distributed computing:	Spark MLlib
Project Management tools:	JIRA, Confluence, Mattermost, Agile
Deep learning toold:	Keras