Kaveri Anuranjana

Personal Data

LinkedIn:	kaveri-anuranjana
Github:	erzaliator
Website:	kaveri.ml
Phone:	+91 8096738698
email:	anuranjana25@gmail.com
Google Scholar:	kaveri-anuranjana

I investigate training methodologies to improve low resource discourse relation classification. My research has contributed tools and resources for Hindi - a language spoken by millions but researched by a few. I have worked on search, semantic similarity and question answering.

Work Experience

Apr 2021- Present | Ph.D. Scholar at Saarland University,
Discourse relation classification
The PhD project entailed developing multilingual classifiers for discourse relation classification (DISRPT task). Developed multi-task classifiers to exploit efficient dataset pairs based on language and label distribution similarities improving classification especially for low resource datasets. Investigating CCR features for improving discourse relation pair representations.

Jul 2019-Apr 2021 | Data Scientist I at Microsoft India (R&D) Pvt. Ltd,
Knowledge Mining and Inference Team
Modelled user search query ranking using click logs. Implemented files grammar which lets users construct NL based queries resulting in 4% increase in coverage for Search in OWA. Prototyped Scrum Update Extractor, a model that extracts user updates and commitments spoken in a video. Worked on developing lightweight, low latency classifiers for automated emails for Outlook and maintaining the datapipeline for user signals like emails. Migrating M2H datapipeline to Azure Data Factory.

Jan-May 2018 | Internship at Goldman Sachs Services Private Limited - India,
Data Sciences Project in Finance and Risk Data Science Department
Analyzed customer trading patterns and company finance activity to predict GCLA requirement.

May-Aug 2017 | Internship at IBM Research - India
Deep Student Response Analysis
Conducted experiments for Student Response Analysis using *deep learning*, determining semantic similarity of questions and student responses.

Jan-May 2017 | Teaching Assistant at IIIT-H
Computational Linguistics 2
Conducted tutorials and held TAship responsibilities for the course *Computational Linguistics 2*.

Apr-Oct 2015 | Research Assistant at Language Technologies Research Centre, IIIT-H,
Website Development Team
Developed modules for virtual experiments hosted for Natural Processing Lab *(morphology, stemming, etc.)*

Jul-Nov 2014 | Research Assistant at VLEAD, IIIT-H, *Testing and Development*
Developed web pages and tested and automated experiment modules hosted by Virtual Labs.

Scholarships and Certificates

Recipient of AFBA scholarship from 2012 to 2016.
Four times consecutive recipient of Dean's List Award from 2015 to 2016.

Technical Skills

Basic Knowledge:	PHP, matlab, java, Prolog
Intermediate Knowledge:	python, html, css, javascript, web2py, HDFS, PySpark, mySQL, c, c++, c#, bash scripting

Publications

K Anuranjana*, V A Rao*, R Mamidi. **Analyzing Curriculum Learning for Sentiment Analysis along Task Difficulty, Pacing and Visualization Axes**. 11^{th} WASSA: Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, EACL, 2021
Processing and Information Systems, 2020.
K Anuranjana*, V A Rao*, R Mamidi. **A SentiWordNet Strategy for Curriculum Learning in Sentiment Analysis**. NLDB, 25th International Conference on Natural Language Processing and Information Systems, 2020.
K Anuranjana*, V A Rao*, R Mamidi. **Hindi Question Generation Using Dependency Structures**. 2^{nd} HAI: Workshop on Humanizing AI, IJCAI, 2019.
K Anuranjana*, V A Rao*, R Mamidi. **HindiRC: A Dataset for Reading Comprehension in Hindi**. 20^{th} International Conference on Computational Linguistics and Intelligent Text Processing, CICLing, 2019.

Major Projects

Deep Student Response Analysis: Experimentation and implementation of an STS classifier, a novel, hybrid deep neural network which uses GloVE embeddings.

Non-factoid QA using RNN: Implementation of Recurrent Neural Network with word2vec in python to answer facebook's bAbi QA dataset to surpass the benchmark results.

Malicious urls Detection: A python implementation and a comparative analysis of classifiers using Machine Learning techniques to detect Malicious URLs

Amazon Customer Review Analysis: A python project on sentiment analysis on Amazon product reviews to crawl the data and perform Aspect based Sentiment Analysis.

Wiki search engine: Python based project to index terms and optimize search queries on a large Wikipedia dump.

Mini linux terminal: Implementation of linux commands using C programming language.

Minor Projects

Language Modeling using via n-grams, Tokenizer, Discourse Marker Detection on Stanford dependency trees, Web scraping and Building Word2Vec embeddings of Ramayana text, Toy Neural Network library, Mini SQL Engine, Pacman game, Carrom board game in OpenGL2, AI agent on 9X9 Tic-tac-toe, TCP/UDP File Sharing Protocol, Hybrid concession based Negotiator Agent, Multi Agent Traffic Light Coordinator.

Education

2021-PRESENT	Ph.D. in Computational Linguistics, **Saarland University**, Germany
2013-2021	B.Tech. in Computer Science and Engineering and MS by Research in Computational Linguistics, **IIIT**, Hyderabad Current cgpa: 8.49/10
2012–2013	Passed Senior Secondary School in CBSE,**Kendriya Vidyalaya,** Sector-8, R.K. Puram, New Delhi Percentage: 92.8/100
2010-2011	Passed Higher Secondary School in CBSE,**Kendriya Vidyalaya (AFS),** Jaisalmer Percentage: 10/10