Dheeraj **Mekala**

(+1) 858 729 8312 | dmekala@ucsd.edu | dheeraj7596.github.io/ | dheeraj7596 | dheeraj7596

Education

University of California, San Diego, MASTERS IN COMPUTER SCIENCE, 4.0/4.0

Sept 2019 - Present

SPECIALIZATION IN ARTIFICIAL INTELLIGENCE

Indian Institute of Technology Kanpur, BACHELORS IN COMPUTER SCIENCE AND ENGINEERING, 8.3/10

Jul 2013 - May 2017

Publications

Contextualized Weak Supervision for Text Classification
Dheeraj Mekala, Jingbo Shang
Association for Computational Linguistics, 2020

User bias removal in review score prediction
Rahul Wadbude, Vivek Gupta, Dheeraj Mekala, Harish Karnick
Proceedings of the ACM India Joint International Conference on Data Science and Management of Data, 2018

SCDV: Sparse Composite Document Vectors using soft clustering over distributional representations
Dheeraj Mekala, Vivek Gupta, Bhargavi Paranjape, Harish Karnick
Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, 2017

Work Experience

Sprinklr India Pvt. Ltd., DATA SCIENTIST

Apr 2018 - Jul 2019

- Architected and built Sprinklr AI's visual insights module, being used by over 1200 Sprinklr clients on a daily basis.
- Developed in-house computer vision models for visual sentiment, gender, age, inappropriate content detection in images and videos.
- Built a scalable system capable of running ML models over 500 million messages per day using Tensorflow, Kafka and Elasticsearch.
- Developed a dockerized auto-scaling python-based framework which is deployed in kubernetes for image classification.

Sprinklr India Pvt. Ltd., PRODUCT ENGINEER

Jul 2017 - Apr 2018

- Developed a centralized monitoring environment(Grafana, InfluxDB) which gathers system metrics as well as docker run-time metrics.
- Collaborated with a team of 3 members and developed an end to end pipeline that incorporated DoubleClick tracking in ads.
- Implemented core functionalities to improve the feature of importing, exporting ads which is the primary way, the users create ads.

Microsoft India, MACHINE LEARNING INTERN

May 2016 - Jul 2016

- Built fully automated case routing system in Microsoft Dynamics CRM, which predicts the ideal assignment candidate for a case.
- Built a robust pipeline which connects Microsoft Dynamics CRM and Azure Machine Learning studio.

Key Projects

Metadata Guided Weakly Supervised Text Classification, RESEARCH PROJECT, PROF. JINGBO SHANG, UCSD

Jan 2020 - Now

- Each text document has some metadata attributed to it. For example, in the research papers corpus, author, venue information.
- We aim to leverage this metadata information to improve the weakly supervised classification.

Contextualized Weak Supervision for Text Classification, RESEARCH PROJECT, PROF. JINGBO SHANG, UCSD

Sept 2019 - Dec 2019

- Proposed a novel framework that performs contextualized weakly supervised text classification.
- Our framework identifies the interpretation of seed words, contextualizes corpus and classifies the documents.
- To the best of our knowledge, this is the first work on contextualized weak supervision for text classification.

Portmanteau Generation, STATISTICAL NATURAL LANGUAGE PROCESSING, PROF. NDAPA NAKASHOLE, UCSD

Sept 2019 - Dec 2019

- Converted text generation task into sequence labelling task by incorporating structural constraints.
- Designed BiLSTM-Regression ensemble model to generate portmanteaus and outperformed several existing techniques.
- Awarded Outstanding Project in the class.

Hierarchical Text Classification, RESEARCH PROJECT, IIT KANPUR

Dec 2016 - May 2017

- Designed $O(nlog(n))$ algorithm to find Bayes optimal hierarchical classification over asymmetric and symmetric loss.
- Designed $O(log(n))$ algorithm to find Bayes optimal hierarchical classification over asymmetric loss under logical assumptions.

Skills

Programming Languages	Java, Python, C, C++, Octave, Matlab, Assembly(x86, MIPS), Bash
Databases & Frameworks	MySQL, MongoDB, Elasticsearch, Spring, ELK, LAMP
ML Libraries & Frameworks	scikit-learn, scipy, OpenCV, Tensorflow, Keras, Caffe, Numpy, Pandas
Software & Utilities	AWS Suite, Docker, Kubernetes, Rancher, Kafka, CI/CD, LaTeX, Git

Relevant Courses

Machine Learning	Machine Learning, Statistical Natural Language Processing, Computer Vision, Probabilistic Graphical Models
Systems	Operating Systems, Database Systems, Computer Organization, Computer Architecture, Compiler Design
Theory	Data Structures and Algorithms, Advanced Algorithms, Theory of Computation