# 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

META: Metadata-Empowered Weak Supervision for Text Classification
Dheeraj Mekala, Xinyang Zhang, Jingbo Shang
*Submitted to Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, 2020

Contextualized Weak Supervision for Text Classification
Dheeraj Mekala, Jingbo Shang
*Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020, Online

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

**Learning Symbolic Music Representation**, Deep Learning for Sequences, Prof. David Kriegman, UCSD — *Mar 2020 - June 2020*
- Designed rhythm-conditioned autoencoder model which generates embeddings of polyphonic multi-track music.
- Was able to successfully reconstruct the music while preserving the rhythm of each track.

**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, Convex Optimization |