Megala Sundar Kannan

(908) 693-3290 | | [msk2245@columbia.edu](mailto:msk2245@columbia.edu) | [www.linkedin.com/in/megalakannan](http://www.linkedin.com/in/megalakannan) | <https://github.com/thisismeg>

2N, 184 Claremont Ave, New York, NY 10027 | US National

**---------------------------------------------------------------------- EDUCATION ------------------------------------------------------------------------**

**Columbia University (CU)** | New York, NY *Aug 2018 - May 2020*

Master of Science in Data Science

*Coursework: Machine Learning for Data Science, Algorithms in Data Science, Probability and Statistics, Exploratory Data Analysis & Visualizations, Statistical Inference and Modelling. Applied ML, Personalization theory & application, Applied Deep Learning, Computer Systems for Big Data*

**College of Engineering, Guindy (CEG)** | Chennai, IN *Aug 2014 - Apr 2018*

Bachelor of Engineering in Computer Science

**------------------------------------------------------------ PROFESSIONAL EXPERIENCE---------------------------------------------------------**

**Johnson & Johnson |** New Jersey, USA*June 2019- August 2019*

Data Science Intern

* Analyzed voice of customer data regarding drug products using Natural Language Processing.
* Developed medical ontologies using Linguamatics and word embeddings (Fasttext) techniques to perform semantic querying.
* Built a text analyzer for the voice of customer data using unsupervised clustering models in python.

**Purple Slate** | Chennai, India *May 2017 - June 2017*

Data Science Intern

* Modeled the users’ real time data using clustering techniques.
* Analyzed price movements in the financial and customer markets.

-**--------------------------------------------------------------------------- SKILLS -------------------------------------------------------------------------**

**Tools & Languages:** Python (Numpy, Pandas, Scikitlearn, Matplotlib, Nltk), R, SQL, PySpark, TensorFlow, Keras, D3.js

**Hard skills:** Data wrangling, Machine learning techniques, Probability, Statistical modelling, Data Visualization, Linear Algebra, Deep Learning techniques, Natural Language Processing.

**-------------------------------------------------------------PROJECTS & PUBLICATIONS ---------------------------------------------------------**

**News Data Analytics |** Capstone Project, Microsoft Research *Jan 2019 – Present*

* Create a platform to analyze news articles and events.
* Topic matching of articles with Wikipedia pages using BERT for community detection.
* Tracking news events across time by creating fingerprints for events.

**Detecting Cancer Metastases in Gigapixel Pathology Images |** Applied Deep Learning, CU*Sep 2019 - Dec 2019*

* Detected tumor cells from pathology images using image segmentation and classification.
* Constructed a convolutional neural network model using tensorflow and keras frameworks.
* Designed evaluation metrics to diagnose the presence of cancer in the cells.

**Recommendation System on Yelp Dataset** | Personalization Theory, CU *Oct 2019 - Dec 2019*

* Built production grade recommendation systems on the yelp dataset for various businesses.
* Predicted ratings on active users using collaborative filtering, non-negative matrix factorization.
* Designed a ‘wide and deep’ learning model for user recommendation.

**Context based Hashtag Recommendation System** | Software Development Lab, CEG*Jun 2017 - Dec 2017*

* Performed topic modelling on twitter posts using machine learning techniques like Latent Dirichlet Allocation.
* Evaluated the results using topic coherence to generate most relevant hashtags.

**Emotion based music player for Android |** 17th IEEE Symposium on Signal Processing and Information *Dec 2017*

Technology, Bilbao, Spain

* Imparted a machine learning approach to perform facial emotion analysis and digital signal processing on audio signals.
* Categorized songs into various emotions by extracting midterm features to compute their valence and arousal values.
* Applied regression through Support Vector Machines was used to train the model on these audio features and a Valence-Arousal coordinate plane was defined to segregate the emotions.