

# Manoj Kumar

☎ (201-899-0725) | ✉ mks542@nyu.edu | 💻 MechCoder | 🐦 MechCoderr

## Current Positon

---

### Google Brain

RESEARCH RESIDENT

Mountain View

July 2017

## Education

---

### New York University

M.S. IN DATA SCIENCE - 3.87/4.0

New York City

Sept 2015 - May 2017

### Birla Institute of Technology and Science

B.E. IN MECHANICAL ENGINEERING - 8.62/10.0

Goa, India

Aug 2011 - May 2015

## Open-source development

---

### Scikit-garden

CO-CREATOR

April 2017 - Present

- Support for Mondrian Forests and Quantile Regression Forests with a scikit-learn API

### Scikit-optimize

CO-CREATOR

June 2016 - Present

- Supports Gaussian Process and Random Forest search for sequential model-based optimisation

### Scikit-learn

CORE DEVELOPER

2013 - Present

- Co-mentor – Google Summer of Code 2016, Cython-fused types support
- Student – Google Summer of Code 2014, Improved speed and memory of linear models

### Apache Spark

GOOGLE SUMMER OF CODE 2015, STUDENT

Remote

May 2015 - Aug 2015

- Improved parity of MLlib's PySpark API

### SymPy

GOOGLE SUMMER OF CODE 2013, STUDENT

Remote

May 2013 - Aug 2013

- Improved symbolic differential equation solving support

## Experience

---

### IBM Spark Technology Center

SOFTWARE ENGINEERING INTERN

San Francisco

June 2016 - Sept 2016

- Improved support for distributed benchmarking and speed of MLlib's Random Forest implementation upto 10%
- Wrote a Python-API for Apache SystemML's MLContext using Py4j

### Telecom ParisTech

RESEARCH ENGINEER, SCIKIT-LEARN

Paris

July 2014 - Dec 2014

- Implemented Birch, an online clustering algorithm as an alternative to KMeans
- Added support for Multinomial Logistic Regression

## Projects

---

## **EmineT: A variational LSTM to generate rap**

INFERENCE AND REPRESENTATION

*Oct 2016 - Dec 2016*

- Used a variational LSTM to generate rap-music with lyrics as additional input

## **Reading Comprehension**

NATURAL LANGUAGE UNDERSTANDING WITH DEEP LEARNING

*Oct 2016 - Dec 2016*

- Implemented variants of an attention-based Reader to do question answering on the Who-did-What dataset in tensorflow

## **Taxi-News**

BIG DATA

*Feb 2016 - May 2016*

- Analyzed the effect of taxi-related news events such as Uber-ban on yellow cab ridership

## **HitOrFlop**

BUSINESS ANALYTICS

*Oct 2015 - Dec 2015*

- Find the features most determinant of a movie's success using pre-released data mined from iMDB