

# KUNAL VERMA

☎ (778)–917–6940 • ✉ [vermak1996@outlook.com](mailto:vermak1996@outlook.com) • 🔗 <https://www.linkedin.com/in/vkunal1996> •  
🌐 <https://www.github.com/vkunal1996> • 🌐 <https://www.vermakunal.com>

## EMPLOYMENT

---

### Graduate Research Assistant

Vancouver, CA

*Natural Language Processing Lab at The University of British Columbia*

*Jun'20 — Aug'20*

- Collected **1 Billion Tweets** of **1 Million users** from **268 countries** in **104** different language using Twitter Python API.
- Inspected the inter-state mobility of the users in the US between January'20 to May'20
- Investigated the Rank change of the **top 40 most visited domains** in 2020 comparing to 2019
- Predicted the undefined languages using **langid**, **langdetect**, and **fasttext** models increasing the total number of detected languages from 67 to 100+ languages.
- Operated **AraNet** on **1 Billion Arabic Tweets** to predicts sentiments, dialect, irony and age-gender of the users.

## EDUCATION

---

### University of British Columbia

Vancouver, CA

*Master of Engineering in Electrical and Computer Engineering*

*Sep'19 — Jun'21*

### Guru Nanak Dev University

Jalandhar, IN

*Bachelor of Technology in Computer Science and Engineering*

*Jun'15 — Jun'19*

## SKILLS

---

**Technical Skills:** Java, Python, JavaScript, C++, HTML, CSS, BootStrap, L<sup>A</sup>T<sub>E</sub>X

**Databases:** MySQL, MongoDB, firebase

**Framework and Tools:** NodeJS, ReactJS, NumPy, Pandas, Flask, Spark(Python), Hibernate, Spring, Struts, JSF, JSP, git, Heroku

**Interests:** Machine Learning, Natural Language Processing, Cloud Computing

## PROJECTS

---

### Mega-COV: A Billion-Scale Dataset of 100+ Languages for COVID-19

- Research Paper accepted at EACL2021 conference.
- Generated a **Billion Scale** dataset to help researchers for studying phenomena related to pandemics.
- Collected more than **1 Billion Tweets** from 268 countries using Tweeter Python API.
- Researched and Examined the **top 20 languages** of the tweets.
- Coordinated to develop a BERT-Base Multilingual Case model for detecting tweets related to COVID-19 with a significant accuracy of 94%.

### Client-Server Chat Application

- Developed a client-server chat application using JavaScript, NoSQL, and NodeJS.
- Examined different cross scripting attacks on the websites and sanitizing process to avoid such attacks.
- Coded the functionality of creating groups, fetching messages from the database server, and connecting clients to the server in order to send and receive messages.

### Emotional Polarity Classifier

- Developed a web based application using flask and python to predict the sentiments of the feedback given by customers to the airline services.
- Deployed and Evaluates different classical and deep learning models for the prediction process.
- Deployed the whole web application and models on the heroku platform.