

# Sumukh S

✉ s.sumukh11@gmail.com    🌐 ssumukh.github.io    📄 github: ssumukh    in LinkedIn    🎓 GScholar    ☎ +91 94824 19972

---

## EDUCATION

### International Institute of Information Technology, Hyderabad (IIIT-H)

B Tech (with Honors) + MS in Computer Science

Integrated dual degree program, with specialization in Machine Learning and NLP

Thesis advisor: Professor Manish Shrivastava, Natural Language Processing(NLP) lab, LTRC

---

Hyderabad, India  
(Expected) June 2021

## WORK EXPERIENCE

### LinkedIn

Data Science Intern

Bangalore, India  
May 2020 - Aug 2020

- Flagship Data Science team: To start recommending events through *My Network tab* from the time of creation, developed models using unsupervised ML and analysed the results. Experimented with GSDMM & LDA + word embeddings.
- Drove a 14% lift in invite acceptance rate. Preliminary analysis with LinkedIn's *Interest Graph* was also done for 100M MAU.
- Set up a UMP flow to calculate CTR of event live video notifications on a daily basis and analysed the results.
- Used: *Python, scikit-learn, gensim, Hive/SQL*

### Indian School of Business (ISB)

Software Engineer (ML) Intern

Hyderabad, India  
Aug 2019 - Nov 2019

- To automate the loan approval process and enable banks to analyse small transactions, built MVPs for the pilot tests using Django, MySQL & Docker, and developed predictive models using Statistical Machine Learning (ML) algorithms.
- Results from pilot test of loan approval automation showed model was better than human agents, led to a publication([link](#))
- Used: *Python, Docker, Django, MySQL/SQL, scikit-learn, XGBoost*

### Onward Assist

Data Science Intern

Hyderabad, India  
May 2019 - Jun 2019

- Worked on identification of cancer in liver tissue WSIs and segmentation of affected areas using Deep Learning algorithms for Computer Vision. Used U-Net FCN architecture to achieve a Jaccard Index of 0.654 and 0.647 for the tasks.
- Used: *Python, PyTorch, scikit-learn*

### MarketFront Software Solutions

Software Engineer Intern

Hyderabad, India  
Aug 2017 - Nov 2017

- To help users fetch and update details of their products and the track their inventory across different platforms, I led the group of 3 to build an Android Application, Webapp and a RESTful API. Custom QR codes could also be generated.
- Used: *Java [Android app development], Django, MySQL, Python, ReactJS/Javascript, HTML5, CSS3*

### IIIT Hyderabad

Teaching Assistant

Hyderabad, India  
Spring 2020, Monsoon 2020

- Head TA for Advanced NLP course (*Graduate level - covers Deep learning + NLP*) in Spring 2020.
  - Graduate TA for Data and Applications course (*200+ students*) in Monsoon 2020.
- 

## PROJECTS (selected)

### Search Engine for Wikipedia

Python ([link](#))

- Used BSBI to create an inverted index of the entire WikiPedia dump to query on & retrieve top results(relevance ranking)

### Custom Language Compiler

C++, STL ([link](#))

- Developed a fully functional front-end of the compiler to generate LLVM IR for a custom C-like programming language

### HTTPS Multithreaded Proxy Server with Cache

Python ([link](#))

- Implemented a proxy server that can handle multiple clients at a time (multithreading) with server-side caching.

### Mini-Dropbox: P2P File Sync

C++ ([link](#))

- Application Level program to keep two local/remote directories synced, similar to Dropbox using sockets.

### SQL Engine

Python ([link](#))

- Built a small SQL engine with support for basic queries, joins and aggregate functions with exhaustive error handling.
- 

## RESEARCH EXPERIENCE & PUBLICATIONS

### Natural Language Processing (NLP) Lab, IIIT Hyderabad

Research Assistant

Hyderabad, India  
May. 2018 - Present

- **Publication:** Detection and Annotation of Events in Kannada at LREC 2020 (Workshop), Marseille, France ([Paper Link](#))
  - Under the guidance of Professor Manish Shrivastava, working on information extraction from unstructured data.
  - Previously: Multimodal representation for Visual Question Answering & Machine Translation for low resource languages.
- 

## TECHNICAL SKILLS & RELEVANT COURSEWORK

**Programming Languages:** Python, C, C++, Javascript, SQL, MATLAB, Java, Bash

**Libraries and Tools:** Docker, PyTorch, scikit-learn, NLTK, OpenGL, Android Studio, Git, Flask, HTML,  $\LaTeX$ , Linux♥

**Relevant Coursework:** Software Engineering, Database Systems, Computer Networks, Compilers, Distributed Systems, Algorithms, Data Structures, Intro to ML, Natural Language Processing(1 & 2), Optimization Methods, Multivariate Analysis