Curriculum Vitae

Manish Kumar

chemicbook.com

Email : sihagmanish36@gmail.com

Mobile : +91 9167856885

Summary

A data science professional at the intersection of natural sciences and technology aspiring to develop AI solutions to further the application of in-silico modeling in accelerating Biopharma R&D

Education

Indian Institute of Technology, Bombay   **[July'14 – July'18]**

Bachelor of Technology in Chemical Engineering

Technical Skills

Programming languages	Python, SQL, JavaScript, HTML, CSS, GraphQL, Bash	3 years+
Familiar with	R, Java, C, C++, Octave, MATLAB, NWChem	1-3 projects
Tools & OS	Kubernetes, Postgres, Neo4j, Docker, Linux, Git, Jupyter, Airflow, VMs	3 years+
Libraries	Tensorflow, Keras, RdKit, ReactJS, Sklearn, Pandas, NumPy	3 years+
Others	Data Science, Machine Learning, Deep Learning, Neural Networks, NLP, RNN	3 years+

Experience

Data Scientist, *Aganitha Cognitive Solutions, Hyderabad*   **[Sep'18 – Present]**

Drug Solubility Prediction to improve drug developability assessment

- Trained Machine Learning models to classify drug molecules into low, medium, and high solubilities; The models achieved **78%** accuracy and are currently **helping real-world scientists** in their day-to-day experiments
- Trained deep learning models to predict quantum descriptors of chemical molecules **reducing** its computational cost by **99.9%**
- Designed a standardized database schema that can efficiently store experimental lab results and molecular descriptors
- Managed and **trained a team** of four freshers by writing tutorials, blog posts, quizzes, and collating study materials

Suzuki Reaction Yield Prediction to decrease avoidable experimental expenditure in Biopharma R&D

- Trained Graph Neural Networks to identify low yielding reactions that helped the client in a **30% reduction** of lab expenditure
- Developed a Jupyter notebook-based app that helps scientists in visualizing chemical reactions and molecular structures
- Leveraged NLP techniques to extract and curate a database of chemical reactions from millions of published patents
- Built an **Airflow** based framework to visually monitor the progress of the data pre-processing ETL pipeline in real-time

Research Publication Browser to enable a faster intelligence solution for R&D scientists

- Created a hub that can search the internet and automatically collate relevant research related to any disease
- Built a ReactJS based website with neo4j as the backend that can **efficiently search** millions of research papers in real-time
- Developed a **master data management** pipeline for disambiguating organization names, author names, and their affiliations

Email Request Classification for streamlining customer care operations

- Developed RNN models to classify customer inquiries into 70+ categories **saving six hours/day** spent in manual classification
- Improved customer experience by automatically assigning ticket numbers to their queries and reducing the waiting time from **two days to five minutes**

In-Silico Antibody Engineering to identify the structure of an antibody

- Created a pipeline that can combine both Homology Modelling and deep learning approach to predict the structure of an antibody, a step towards saving months of experimentation in the lab
- Created data analysis programs to identify low and highly viscous monoclonal antibodies

Machine Learning Intern, *Eli Research India, Faridabad*   **[May'17 – July'17]**

Chatbot, *a QA chatbot to answer common and repetitive queries by customers*

- Carried experiments using RNNs and GRUs to provide word prompt when a user starts typing
- Helped the company in reducing the number of customer calls by **10%** by setting up a QA chatbot

RELEVANT COURSES

- Deep Learning Specialization, *deeplearning.ai, Coursera*   **Aug'17 – Feb'18**
- Natural Language Processing, *Stanford University, YouTube*   **May'17 – Jun'17**
- Machine Learning, *Stanford University, Coursera*   **Mar'17 – Jun'17**
- The Data Analytics Edge, *MITx Courseware, edX*   **Jan'17 – Feb'17**
- Algorithms, *University of California, San Diego, Coursera*   **Dec'16 – Jan'17**
- R Programming, *Johns Hopkins University, Coursera*   **Aug'16 – Sep'16**
- Introduction to Computer Science, *Harvard, edX*   **Jul'16 – Nov'16**
- Introduction to C++, *IIT Bombay, On-Campus*   **Aug'14 – Nov'14**