**Vivek MODI**
Third Year Undergraduate
Discipline of Computer Science and Engineering
vivek.modi@iitgn.ac.in
+917383080241

- I am a highly motivated and proficient Computer Science graduate student with a strong foundation and experience in machine learning, data science, and programming. Currently, I am working as a Software Engineer- Machine Learning at Optimoz Inc. I completed my Masters in Computer Science from Rutgers University. I completed my undergrad in Computer Science from Indian Institute of Technology, Gandhinagar(IIT). I secured an All India Rank of 701 out of 1.3 million students in the JEE ADVANCED 2018. My educational journey is complemented by practical experiences, including a successful summer internship at Capgemini, where I developed an email prioritization add-on using sentiment analysis and machine learning. My research paper, "ComicBot," demonstrates my expertise in NLP and creative problem-solving. My academic and project-based achievements, along with my proficiency in Python, data analysis, MLops and an array of relevant tools, make me a versatile candidate for roles in machine learning, data science, and beyond. With a passion for clean code, continuous improvement, and a collaborative mindset, I'm poised to contribute effectively to any dynamic team.

## Education

| Degree | Institution | Year |
|---|---|---|
| MS | Rutgers University | 2022 - 2024 |
| B.Tech | IIT Gandhinagar | 2018 - 2022 |
| Class XII | LBS Convent School, Kota | 2017-2018 |
| Class X | Divine Child School, Mehsana | 2015-2016 |

## Projects

- As a Machine Learning Engineer (GenAI/ML) at Optimoz, I developed and deployed advanced LLM-based solutions for Optalk.ai, focusing on domain-specific fine-tuning, RAG pipelines, and citation extraction from scanned PDFs using OCR and NLP techniques. I improved clinical data extraction accuracy and built scalable ML pipelines that significantly enhanced system performance. Notably, I reduced chat response latency by 85% through backend optimizations and streamlined model workflows. My work bridged research and deployment in real-world healthcare and enterprise NLP applications.
- As a Software Engineer (GenAI/ML) at Optimoz, based in Rockville, MD, I have been pivotal in developing and implementing a sophisticated Large Language Model for Optalk.ai, significantly enhancing its AI capabilities. My role involved engineering robust and scalable backend systems using Flask and Kafka, integrating and optimizing advanced Generative AI models for language generation tasks, and leading the development of a streamlined CI-CD pipeline to ensure efficient and reliable software updates and deployments.

- **Software Engineer(GenAI/ML), Optimoz, Rockville, MD [May, 2024-Present]**
  - Developed and implemented a Large Language Model for **Optalk.ai**, enhancing the platform's AI capabilities.
  - Built backend systems using **Flask** and **Kafka**, ensuring robust and scalable performance.
  - Integrated and optimized Generative AI models for **Language Generation** task.
  - Streamlined deployment through a **CI-CD** pipeline, ensuring efficient and reliable updates.
- **Graduate Research Assistant, Machine Learning & Bioinformatics Lab, Rutgers University [Jan, 2023 - May,2024]**
  - Built a comprehensive deep Learning pipeline for the identification of Human Activity.
  - Implemented a ML learning-based pipeline to identify protein subcellular sequences working under the guidance of Dr. Iman Dehzangi.

- As a Graduate Research Assistant at the Machine Learning & Bioinformatics Lab, Rutgers University, I led innovative projects in Human Activity Recognition (HAR) and protein subcellular localization prediction. I utilized CNN, LSTM, Multimodal Transformer, and Action Transformer models to improve detection accuracies, achieving 88.4% accuracy with the Action Transformer. I developed a machine learning pipeline under Dr. Iman Dehzangi's guidance to predict protein subcellular localization in Gram-positive bacteria, achieving 87.237% accuracy by analyzing protein sequence features. My research significantly advanced bioinformatics, contributing to understanding complex human activities and cellular processes.

- **Summer Internship at Capgemini, Gandhinagar under Mr. Ashish Buch.**
  - Developed an **add-on in outlook** for prioritizing the mails based on importance.
  - Implemented sentiment analysis using two distinct machine learning models.
  - Created the Priority Mailbox and Sentiment Analysis COM-addin for Microsoft Office Outlook Desktop Application.

- Utilized Django for web development, employed various machine learning algorithms, and managed data using an SQLite database.
- Utilized multiple parameters such as the recipient's first interaction with the email, the time taken to read the email, and other factors to determine email prioritization.
- 
- Developed: Designed and created an Outlook add-on geared towards mail prioritization based on criticality.
- Implemented: Successfully integrated Sentiment Analysis into the system, utilizing two distinct models.
- Created: Developed a Priority Mailbox and Sentiment Analysis COM-addin for Microsoft Office Outlook Desktop Application.
- Utilized: Employed Django, machine learning algorithms, and an SQLite database as the core technologies.
- Enhanced: Improved mail prioritization by considering multiple parameters, including initial click time and mail read duration.
- Contributed: Played a pivotal role in optimizing mail organization and enhancing productivity within the Outlook platform during the internship.

- During my internship as a Machine Learning Intern at Capgemini in Gandhinagar, working under the guidance of Mr. Ashish Buch, I had the opportunity to contribute to an exciting project. My primary focus was on developing a Priority Mailbox and Sentiment Analysis COM-addin for Microsoft Office Outlook. This involved leveraging a combination of technologies, including Django, machine learning algorithms with two distinct models, and a SQLite database.
  One of the key achievements during my internship was the significant improvement in mail prioritization. We accomplished this by considering multiple parameters, such as initial click time and mail read duration. By doing so, we played a pivotal role in optimizing mail organization within the Outlook platform, ultimately leading to enhanced productivity. This internship experience was not only rewarding but also provided valuable insights into the practical application of machine learning in real-world scenarios.
- **Protein Subcellular Localization Prediction Using Machine Learning [Sep,2023-Jan,2024]**
  - Utilized machine learning techniques to predict **protein subcellular localization** in Gram-positive bacteria, focusing on four cellular locations.
  - Extracted and analyzed protein sequence features like Occurrence and Composition, attributes like structural **hydrophobicity** and **polarizability**, and prepared labeled datasets for multi-class classification tasks.
  - Enhanced understanding of protein functionality in cellular processes, contributing to advancements in bioinformatics and computational biology.
- **Machine Learning Engineer, Coder Edge Technology, Mehsana, India [Jan,2022-Jul,2022]**
  - Developed an **Automatic License Plate Recognition** system, achieving **96.53% accuracy** under constant lighting, using Python, **OpenCV**, and **TensorFlow**.
  - Optimized machine learning models for real-time recognition of diverse license plate formats in live surveillance footage.
  - Collaborated with Regional Transport Office(**RTO),Mehsana** on system testing and integration, enhancing traffic management and security at industrial sites.
  - Dev pandya - +919724348749

As a Machine Learning Engineer at Coder Edge Technology in Mehsana, India, I developed and optimized an Automatic License Plate Recognition system that achieved 96.53% accuracy. Utilizing Python, OpenCV, and TensorFlow, I tailored machine learning models for real-time recognition of various license plate formats from live surveillance footage. My collaboration with the Regional Transport Office (RTO) in Mehsana focused on rigorous system testing and integration, significantly enhancing traffic management and security at local industrial sites.

- **Captcha-Breaker to recognize captcha using Machine learning models. [Jan,2021-May,2021]**
  - Explored various strategies for deciphering captchas across multiple classifications.
  - Created a machine learning algorithm specifically designed to tackle **anti-recognition** captcha challenges.
  - Addressed **anti-segmentation** captchas by applying advanced deep learning techniques.
- **NLP project "ComicBot: ChatBot Generating Jokes along with GIF" under prof. Mayank Singh.**
  - **First** Research paper on Generating Jokes along with Gif with **knowBERT** model.
  - Used **Sarcasm detection and Emotional Classifier** for mapping with **most related GIFs**.
  - Proposed a **new Dataset "EmotionGIF"** for describing labels for GIFs.
  - Proposed a **Style transfer** approach for generating Jokes.
- **Human Activity Recognition Using Machine Learning under Prof. Iman Dehzangi [Nov,2022-May,2023]**
  - Employed CNN, LSTM, **Multimodal** Transformer, and **Action Transformer** in a HAR project.
  - Implemented and fine-tuned machine learning models with diverse datasets like mPOSE-21, and UCI HAR.

- Achieved notable performance metrics: Multimodal Transformer (**84.05%** F1-score), Action Transformer **(88.4%** accuracy).
- Contributed to advancing Human Activity Recognition through machine learning.

- **Mini-StackOverflow using flask and React.js under prof. Sameer Kulkarni**
  - Developed a server client-based StackOverflow using flask and React.js
  - Efficient and smooth implementation of several similar features as StackOverflow.
  - Used MySql Database to manage the to and fro of data linked with the flask server.
- **Captcha-Breaker to recognize captcha using Machine learning models**
  - Identified different approaches to **solve captcha** of different categories.
  - Developed a Machine learning model to identify Anti-recognition captchas.
  - Worked to solve **anti-segmentation** captchas using Deep learning models.
- **MacroHard: E-Commerce Website for Technical and Non-Technical Projects**
  - Facilitated user engagement with features like comments, image uploads,and many other serivces.
  - Designed the website using Node.js and React for a modern interface.
  - Utilized a lot various AWS services for cloud-based hosting and operations as well as database management.
  - Delivered a robust global online marketplace for project trading.
- **Mini-Twitter with MiniNet under prof. Sameer Kulkarni**
  - Developed a server-client based twitter using flask.
  - Used **Mininet** approach to effectively measure the inbound processes.
  - Pandas dataset and implemented it on the platform
- **Multi-threaded web crawler with effective web ranking**
  - Implemented a **multi-threaded web downloader** to download multiple sites at a same time
  - Used PageRank algorithm to rank effectively **rank the pages**.
  - Used HTML downloader, Link extractor, domain extractor, and other models.
- **Car Price prediction using Machine Learning.**
  - Analyzed price of the car using **different regressor models**.
  - **96.38%** accuracy using different regressor models.
  - Helped **Verzeo** organization with the **best Car Prediction algorithm**.
- **Data Analysis of Average User Rating in App Store.**
  - Predict the **Genre of apps** that has highest Average User Rating.
  - Inferences of the app to get **highest Average User Rating.**
  - **85% possibility** to get Average user rating greater than 4 for entertainment apps.
- **Handwriting Text Recognition with Metis(Coding) Club.**
  - Accurately recognize text from a given sentence using **Neural Network.**
  - Used **IAM dataset** to test the Text Recognition Project.
  - Implemented using **Tensorflow** with accuracy of **97.16%**.
- **Implemented Firefly Algorithm from scratch for Optimization. [Sep,2022-Dec,2022]**
  - Proficiently implemented the Firefly Algorithm, a **nature-inspired** optimization, for complex problem solving.
  - Designed a custom algorithm mimicking firefly behaviors, enabling efficient exploration of solution spaces.
  - Applied Firefly Algorithm to **real-world optimization** tasks, showcasing its versatility and effectiveness.
  - Attained impressive optimization results, emphasizing practicality and robustness in diverse applications.
- **Digit Recognition using MNIST dataset.**
  - **75% Accuracy** for recognizing digits from 0-9.
  - Developed a good command over **Tensorflow** and its libraries.
- **Cache controller in FPGA using Xilinx Vivado.**
  - Implemented a **Cache controller** for two layers of **Cache Memory** on a **Basys-3 FPGA**.
  - **Performs Read and write operations, L1 - Direct mapping and L2 - four Way Set Associative.**
  - Adopted **Write Back, No Write Allocate, Least Recent Used Algorithm** polices for accessing caches.
- **Harnessing in-pipe energy from tap and grey water at IIT Bombay INTER IIT Tech meet.**
  - **Represented IITGN** at Inter IIT Tech meet for **Campus Sustainability Challenge**.
  - Experimented the pressure in the **water and drainage pipes** of IITGN.
  - Developed a method to generate electricity from water drained from the washroom.
  - Secured **8th position** for IITGN at **Inter IIT Tech meet**.
- **Tic tac toe game, under Prof. Neeldhara Misra.**
  - Created a game in python with the help of **pygame**.
  - Developed new ways to build a game using **hard-core coding** in python.
- **Extracted information Variation of concentration in the Batch reactor under Prof. Indranath sen Gupta.**

- Solved the variation of concentration using rate laws for multiple reactions.
- **Lead** a group of 6 students and **found an equation** to solve Variation of concentration

- **Manufacture and designed a Quadracycle from 2 cycles under prof. Madhu Vadali**.
  - Designing of a Quadracycle and manufacture it.
  - Learned some basic technics of joined two cycles using welding.
  - Developed a fully functional Quadracycle and tested it.

## Positions of Responsibility

- **Organiser, Jashn'18 IITGN**
  - Organised the event **Slammer** in the inter-college cultural fest**.**
  - **More than 100** Participation in a single event.
- **Organiser, CCL'19 IITGN**
  - Organised the only **Intra college Cricket tournament**.
  - Managed overwhelming participation for the tournament with approximately **200** people.

## Achievements

- **Summer Internship** at **Capgemini, Gandhinagar**
- **Certificate** from Verzeo for **Machine Learning** under Prof. Harsha V.
- **Machine Learning Certificate** course by Stanford Institute by Andrew Ng. in Coursera. (Ongoing)
- Ranked **8th** in INTER IIT Tech meet at IIT Bombay for Campus Sustainability Challenge.
- Secured **4466** rank in JEE Advance 2018.
- Secured **14054** rank in JEE Mains 2018.
- Explored and managed a tour in 12 states for 42 days as an **Explorer fellowship** program.

## Skill Summary

- **Languages:** Python, C, HTML, CSS, MySQL, Javascript
- **Tools:** Autodesk Inventor Professional, Latex, LTSpice, Matplotlib, Tensorflow, Vivado

## Extra-Curricular Activities

- Plays keyboard and participated in the Battle of Bands at IITGN.
- Played cricket for a team in CCL.

vivekvm84001@gmail.com +12679285072 https://github.com/viper-vm

# Vivek MODI

Graduate Student
Master's in Computer Science

vivekvm84001@gmail.com
+12679285072

Twitter sentiment analysis

Graduate Student in Computer Science with 2 years of professional experience in the cloud software industry with strong foundations in Networks, Algorithms, Mathematics, programming, and Databases, seeking full-time opportunities in the Technology domain to enhance my abilities and contribute to my organization.

MS Computer Science student at Rutgers University and BTech graduate from IIT Gandhinagar. Specialized in Machine learning, NLP, and full-stack development with Django, flutter, AWS, React and Node.js. Aspiring to leverage my educational background and project experience towards innovative solutions in technology, aiming for roles that challenge and enhance my skills while contributing to the advancement of the field.

Describe your experience with data science and python

I've been immersed in the world of data science and Python, right from my academic journey at IIT Gandhinagar to my masters at Rutgers University. My knack for Python was further polished through hands-on projects and professional stints, such as my role at Capgemini where I leveraged Python to devise ML algorithms for email prioritization, and my research work under Dr. Iman Dehzangi at Rutgers. These experiences have not just enhanced my technical prowess but also enriched my understanding of applying machine learning to real-world problems. Be it developing deep learning pipelines or tackling NLP challenges, my journey has been a blend of learning, application, and innovation, all through the lens of Python and data science.

Describe your experience applying data science techniques in the insurance industry
As my experience primarily spans sectors like bioinformatics, machine learning research, and software development, I haven't directly worked in the insurance industry. However, the skills I've developed, such as deep learning, machine learning, and natural language processing, are highly transferable and can be adeptly applied to insurance. For instance, the techniques I used in projects like sentiment analysis for email prioritization could easily translate to analyzing customer feedback or claims data in insurance, helping to streamline processes and improve customer satisfaction. My strong foundation in Python and data science would enable me to quickly adapt and contribute to data-driven solutions in the insurance domain, despite my lack of direct experience in the industry.

If hired, when are you available to begin?
I'm genuinely excited about the possibility of joining your team. I'm currently completing my final semester at Rutgers University, with graduation scheduled for this May. Therefore, I'd be available to start working from May onwards, after my graduation. I'm looking forward to the opportunity to bring my skills and enthusiasm to your organization and contribute to impactful projects.

What are your salary expectations?
I'm very much looking forward to the possibility of working with your team. I trust your organization to offer a fair compensation that aligns with the industry standards and reflects the role's responsibilities and my experience. A figure around $100,000 would be wonderful, but I'm open to discussing whatever you believe is appropriate and fair.

What is your most impressive work accomplishment?
I have worked in the field of Machine Learning for quite some time now but during recent time, i have been engaged a lot in the development of Large Language Models from scratch for an industry. It was fascinating for me to build a LLM from scratch with compliance to the industry's requirements. I am working with the LLM for quite some time for my projects but it was different feeling as well as accomplishment to complete the task for a firm.