

HARSH SAGAR
harsh.sagar2107@gmail.com
6352010302

2nd February, 2024

Boston Consulting Group,
India

Hello,

I am Harsh Sagar, 5th year Biological Sciences student at IIT Madras. I wish to apply for Data Scientist position at Akaike Technologies. I had worked on Detection of Milk Adulterant using ML Techniques project for Summer Research Internship, IIT Bhubaneswar. Analysed milk properties such as foamability, foam stability, foam height etc as the demand for Cappuccino foams has increased. Used Grid search cv for hyperparameter tuning on Random forest classifier, which increased test accuracy to 88.1% from 81.7% also increased roc auc score to 0.953771 from 0.946013.

Worked on the classification of the gene expression of endometrial biopsies as a data science intern at Supratech Lab, Gujarat. To select the genes involved in endometrial receptivity to be included in the ERA, the gene ontology terms with a P value less than .025 were considered significant. For the endometrial receptivity transcriptomic signature, two different comparisons were made, the first was the differential expression of R versus P and E and the second was that of P versus E in the ERA. These statistically significant genes were isolated during the Window of Implantation. Used Grid Search CV for hyper parameter tuning of XGBoost model which was used for prediction.

Working on these internships I have developed extensive knowledge and expertise in the field and gained valuable experience in various aspects of research work. Moreover, my previous roles have helped me to develop and significantly improve my attention to detail and accuracy and gave me a chance to see and feel what it's like to be a part of a deadline-driven and fast-paced business environment. In addition to my experience, I can offer your firm excellent analytical skills and important multitasking and time management abilities. Not only they are necessary for the job but also serve as an evidence of a data-driven mindset which I believe I possess.

My proficiency in Python, along with my knowledge of languages such as R and Java, positions me well to tackle the diverse challenges outlined in the job description. I have hands-on experience in constructing robust data pipelines, handling structured and unstructured data for model training, and deploying AI/ML models into production. This hands-on experience is complemented by my understanding of Generative AI pipelines, including orchestration and prompt engineering.

I am particularly excited about the opportunity to leverage my skills in building AI orchestration with frameworks like LangChain, working with vector databases such as Pinecone and Chroma, and utilizing data processing frameworks like Data Bricks and Airflow. My familiarity with API frameworks, including Django and Flask, further contributes to my readiness for this role.

Over and above, I am a good team player also I have ability to compromise, listen to my teammates and always passionate to learn from my seniors. I'm reliable and willing to work above and beyond what is required to get going on any assignment with enthusiasm.