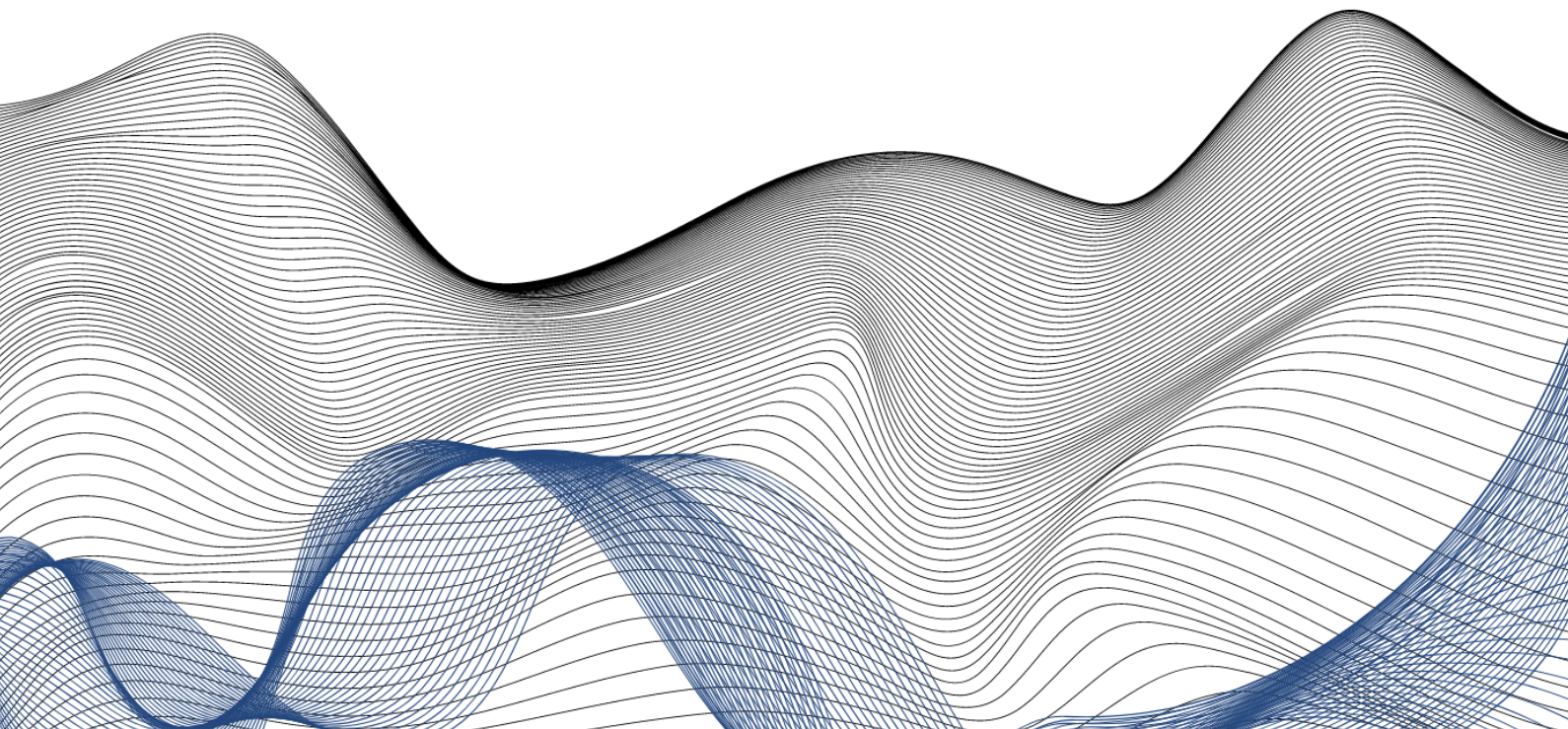




Data Scientist / Machine Learning Engineer

Saigen | Job Description



About us

Saigen is a leading provider of automatic speech recognition (ASR) in Africa, specialising in the development of high-quality, cost-effective speech technologies for a growing list of African languages, including isiZulu, Sesotho, Setswana, Swahili, Twi, Luganda, Amharic, Afrikaans and English, with more in the pipeline. Our core mission is to bridge the language gap and enable organisations to become truly multilingual, whether they're processing customer interactions in a call centre, monitoring the media, or delivering real-time transcriptions.

Saigen is part of the Alphawave group and a shareholder in Callbi, a key player in speech analytics in Africa. We collaborate closely with Callbi and recently formed an exciting new joint initiative, Callbi AI Labs, that focuses on solving specific client business problems using insights derived from speech and other data.

If you're passionate about building cutting edge ASR systems for new languages, or solving business challenges by combining insights from call centre speech with other business data, Saigen could be the perfect fit for you!

Our Team and Environment

We are a dedicated, skilled, and curious team that thrives on innovation and problem-solving. Our startup culture fosters a collaborative and supportive environment where unconventional thinking is encouraged, and diverse approaches to challenges are explored. We value commitment to detail, dedication, and hard work, and we are deeply invested in the personal growth and development of our team.

Our Offices

Our offices are located within Brickfield Canvas in Salt River, offering flexible working hours to promote a healthy work-life balance. On-site, we provide unlimited barista-style coffee and easy access to the cafeteria for ordering food.

Job Description: Data Scientist / Machine Learning Engineer

We are seeking a passionate and driven Data Scientist/ML Engineer to join our team. The job description below outlines the key responsibilities and qualifications for the role; however, we recognise that candidates may not have expertise in every area. What we value most is a genuine enthusiasm for technology, a commitment to continuous learning, and the ability to develop impactful models and build reliable systems. Equally important is being a collaborative team player who prioritises the collective success of the team and the needs of our clients.

As a Machine Learning Engineer at Saigen, your main goal will be to advance our ASR, NLP and AI capabilities. You'll focus on building, optimising, and deploying machine learning models that improve speech recognition accuracy for African languages, with additional opportunities to contribute to our media monitoring, real-time transcription, text-to-speech (TTS) and AI Labs solutions.

Key Responsibilities

1. Speech & Language R&D

- Conduct research on state-of-the-art ASR, NLP, and generative AI methods (including LLMs) to enhance multilingual speech recognition and text-to-speech quality.
- Design experiments and develop novel approaches to overcome linguistic challenges specific to African languages.

2. Model Development & Optimisation

- Design and implement end-to-end ML pipelines—from data collection and labeling to model training, evaluation, and deployment.
- Utilise deep learning frameworks (e.g., PyTorch) and ASR/NLP libraries (e.g., Hugging Face) to advance model accuracy, speed, and scalability.

3. Product Integration

- Collaborate with software engineers and linguists to integrate ASR and TTS modules into Saigen's media monitoring, real-time transcription, and analytics services.
- Ensure seamless model updates and performance monitoring, whether deployed on-premise, in the cloud, or in client environments.

4. Client Engagement & Custom Solutions

- Work with project managers and client success teams to understand each client's unique requirements, from call centres to media outlets.
- Tailor solutions to meet compliance, operational, or experience goals, and showcase how our technology can transform their workflows.

5. Quality Assurance & Improvement

- Define, track, and report KPIs (e.g., word error rate, latency) to maintain high standards for transcription and TTS performance.
- Continuously audit and refine ML models, leveraging feedback from production data and user input.

Qualifications and Skills

● Education

- Bachelor's degree in Engineering (B.Eng) or Computer Science (B.Sc. Compsci) is required.
- A Master's degree in a related field.

● Experience

- At least two years of professional experience in data science, machine learning, or a related field.

● Technical skills

- Excellent Python skills, including experience with essential data science and data visualisation libraries (pandas, numpy, scikit-learn, matplotlib, seaborn)
- Familiarity with various machine learning algorithms and experience with deep learning libraries (PyTorch).
- Proficiency in building and consuming RESTful APIs
- Familiarity with containerisation technologies (e.g., Docker)
- Experience with data manipulation techniques—particularly text data (e.g., regular expressions)—for data cleaning and transformation.
- Proficiency with version control systems (e.g. Git)

● Soft Skills:

- Excellent communication and collaboration skills, with the ability to work effectively in a team-oriented, fast-paced environment.
- A proactive attitude towards learning and staying current with the latest trends and developments in data science and speech analytics.
- A meticulous approach to data processing, analysis, and model development, ensuring accuracy and consistency in all tasks.
- Ability to adapt to new challenges and contribute innovative solutions.
- Ability to work independently and lead a small team.

● Preferred Skills

- Experience with ASR
- Experience with generative AI technologies such as LLMs.
- Experience with text analytics and natural language processing (NLP).

- Familiarity with relational and NoSQL databases.
- Cloud experience (AWS/Azure)

What can you expect from Saigen?

- Competitive salary
- Generous leave policy
- Flexible working hours and hybrid policy
- Opportunities to grow and develop your career
- Offices with lunch and barista-style coffee

How to Apply?

Send your CV to careers@saigen.co.za with the subject Saigen *ML Engineer*. Please feel free to reach out if you have any questions.