Audio Data Analyst Trainee (Transcription & Entity Extraction)

DataKlout (Every Call is a Business Opportunity) is an AI & Cognitive solution that is designed to understand the human voice & speech-based conversations. Our platform analyzes conversations, providing actionable insights, trends, and analysis to help businesses make informed decisions. We are embarking on a new generation of products built on Cognitive & Artificial Intelligence. Join us to lead, learn, and grow your career while contributing to building something truly remarkable.

We are looking for a highly motivated **Audio Data Analyst Trainee** to join our AI/ML Engineering team. The ideal candidate should have an interest in audio processing and transcription, with a strong ability to work on annotating and extracting entities from large audio dataset to assist in AI model training. This is a fantastic opportunity to work in a fast-paced, cutting-edge environment with the potential to grow into a key contributor in AI/ML projects.

Responsibilities:

- Label and annotate audio files with transcriptions and corresponding entities.
- Analyze and process audio data, preparing it for machine learning model training.
- Support the machine learning and data science teams in ensuring high-quality datasets for AI/ML models.
- Use audio editing tools (e.g., Audacity, Adobe Audition) to manipulate and enhance audio files under supervision.
- Follow established guidelines for audio annotation and entity extraction.
- Participate in team meetings and provide feedback to improve transcription and entity extraction models.
- Assist in organizing and maintaining datasets for ongoing AI/ML projects.
- Ensure consistency and accuracy in labeling and entity extraction tasks.
- Work collaboratively with engineers and data scientists to enhance audio data processes.

Soft Skills:

- Excellent verbal, written, and communication skills.
- Strong attention to detail with the ability to handle large datasets accurately.
- Collaborative team player with strong interpersonal skills.
- Self-motivated, flexible, and eager to learn.
- Ability to manage tasks efficiently in a fast-paced environment.

Technical Skills and Experience:

- **Fresher candidates** with a strong interest in audio processing, transcription, and machine learning are welcome.
- Familiarity with audio editing tools (e.g., Audacity, Adobe Audition) is a plus.
- Basic understanding of transcription tasks or experience in text-based data labeling is an advantage.

- Knowledge of Python or other programming languages is required.

- Ability to work with large datasets and ensure data accuracy.

- Basic knowledge of machine learning and AI principles is a plus.

Preferred Qualifications:

- Experience in creating neural networks for voice data is a plus.

- A background or coursework in linguistics, data science, artificial intelligence, or related fields.

- Interest in Natural Language Processing (NLP) or speech recognition technologies.

- Familiarity with NLP tools and libraries is an advantage.

- Experience with data labeling or annotation in a machine learning context is preferred.

Education:

- BTech in Computer Science, Data Science, Engineering, or a related field.