Category	Extracted Detail	Role Relevance	Notes (If Any)
Tech Stack	Python (version 3.11 preferred), R	Must-have	Explicitly required for production-grade ML pipelines and general data science tasks.
Tech Stack	SQL & Big Data querying tools (Spark, PySpark, Hive, Presto)	Must-have	Crucial for handling 'Bharat-scale' data ingestion and querying across billions of rows.
Skills	Predictive modeling, Clustering, and Segmentation	Must-have	Applied for behavioral analysis, consumption forecasting, and user profiling.
Skills	Statistical Modeling & Hypothesis Testing	Must-have	Essential for A/B testing frameworks and experimental design.
GenAI	Generative AI, LLMs, and Foundation Model fine-tuning	Strong Advantage	Focused on 'Agentic AI' using specialized agents and Small Language Models (SLMs).
Audio-Video Analytics	ASR (Automatic Speech Recognition), TTS (Text-to-Speech), STT (Speech-to-Text)	Strong Advantage	Used for media enhancement, metadata generation, and real-time annotation.
Audio-Video Analytics	Computer Vision (OpenCV, YOLO) and Multimodal Learning	Strong Advantage	Applied to identify emotions/objects in video content for 'Moment.AI'.
Use Case	Social Media Analytics (Sentiment Analysis, Theme Detection, NER)	Must-have	Analyzing interactions on Instagram, Facebook, and Twitter to decode viewer engagement.
Intern Responsibility	Building production-grade ML pipelines and modular code	Must-have	Expectation to write clean, efficient code for forecasting and recommendation engines.
Intern Responsibility	Collaborating with cross-functional stakeholders	Must-have	Aligning outcomes with business, content, and engineering teams.
Interview Focus	SQL Proficiency (Difficulty 9/10)	Must-have	Questions focus on window functions, multilevel aggregations, and subqueries.
Interview Focus	Case Studies (Churn Prediction & User Churn Analysis)	Must-have	Candidates must propose data-driven strategies to identify and retain high-risk users.
Tools	Big Data Platforms (Hadoop, Hive, Pig, Cassandra)	Strong Advantage	Preferred experience for handling large-scale datasets.
Tools	Digital Analytics (SiteCatalyst, Google Analytics)	Optional	Listed as a 'plus' in the job description.