VosynVerse HK Limited - ML Evaluation & Data Systems Engineering - Master's/Doctorate-Level Postgraduate Internship

Company Name: VosynVerse HK Limited
Position: ML Evaluation & Data Systems Engineering - Master's/Doctorate-Level Postgraduate Internship
Posting Date: 07 Oct 2026

Job Title	ML Evaluation & Data Systems Engineering - Master's/Doctorate-Level Postgraduate Internship
Job Description:	**About the Role:** We are seeking a motivated and technically capable master's- or doctorate-level recent graduate or final-year student to join our Hong Kong research and development team as an ML Evaluation & Data Systems Engineer. This role focuses on building the systems used to measure, validate, reproduce, and improve the performance of multilingual AI models and applications. You will help create evaluation frameworks, multilingual benchmarks, data pipelines, experiment-tracking systems, audit logs, and reporting tools that allow Vosyn to demonstrate how well its AI systems preserve meaning, context, tone, cultural relevance, and accuracy across languages. This is an important role within our AI development process. You will help ensure that model improvements are measurable, technical experiments are reproducible, data is traceable, and product-quality claims are supported by reliable evidence. **Key Responsibilities:** • Design and develop evaluation frameworks for multilingual large language models and localization systems. • Build automated and human-review workflows for measuring AI output quality. • Establish benchmarks for contextual accuracy, semantic preservation, cultural appropriateness, consistency, latency, hallucination, safety, and overall usability. • Develop multilingual test datasets and evaluation scenarios covering Cantonese, Mandarin, English, and other strategically relevant languages. • Create processes for dataset ingestion, cleaning, validation, annotation, version control, and quality assurance. • Develop data pipelines that support model evaluation, experimentation, reporting, and continuous improvement. • Implement experiment-tracking systems for model versions, prompts, configurations, datasets, outputs, and evaluation results. • Design audit-log structures that connect source content, model configuration, output, reviewer feedback, quality scores, and corrective actions. • Build dashboards and reports that communicate AI performance to technical teams, management, pilot partners, and other stakeholders.

	• Support the design of human-evaluation rubrics, reviewer instructions, quality standards, and disagreement-resolution processes. • Apply statistical methods to compare model versions and determine whether observed improvements are meaningful. • Monitor model and data quality and help identify performance degradation, inconsistencies, bias, or unexpected behaviour. • Collaborate with applied AI engineers to test model improvements and identify areas requiring further development. • Support reproducible model experimentation, validation, and pilot-readiness testing. • Document system architecture, evaluation methodologies, data lineage, technical decisions, results, and recommended improvements. • Maintain records that support product development, responsible AI practices, intellectual property documentation, and future enterprise requirements.
Requirements:	• Currently enrolled in or recently graduated from a master's or doctorate program in Computer Science, Artificial Intelligence, Machine Learning, Data Science, Data Engineering, Statistics, Software Engineering, Information Systems, or a related field. Enrollment in or completion of a master's or doctorate program is mandatory. • Strong programming ability in Python. • Proficiency in SQL and experience working with structured and unstructured datasets. • Strong understanding of machine learning, data pipelines, model evaluation, or analytical methods. • Familiarity with large language models, natural language processing, transformer-based systems, or generative AI. • Experience with model evaluation, experiment design, statistical testing, benchmarking, or human-evaluation methods is an asset. • Familiarity with databases, APIs, cloud environments, data visualization, and dashboard-development tools. • Experience with experiment-tracking or MLOps platforms is an asset. • Familiarity with tools such as MLflow, Weights & Biases, DVC, Airflow, Prefect, or comparable technologies is an asset. • Understanding of data lineage, model observability, data governance, responsible AI, or auditability is an asset. • Experience working with multilingual data or language-based AI systems is an asset. • Professional proficiency in Cantonese or Mandarin is a strong asset. • Strong analytical, organizational, and problem-solving skills. • Excellent attention to detail and a disciplined approach to documentation and version control. • Ability to translate technical findings into clear reports and

	recommendations. • Effective written and verbal communication skills. • Ability to work independently while collaborating with a distributed, multidisciplinary team. • Must be based in Hong Kong, or able to work from Hong Kong, and legally eligible to work in Hong Kong. Exceptional talent often emerges from diverse paths. If you possess profound curiosity, a genuine passion for continuous personal and professional growth, and a strong desire to apply your abilities to create meaningful impact, we encourage you to apply even if your background does not align perfectly with every qualification.

Application method: Please apply through https://careers.vosyn.ai/?job=102