Workshop Title	AI and Data Science Workshop for Educators
Duration	2 Days
Target Audience	AI & Data Science Educators

Description
In today's rapidly evolving technological landscape, understanding the fundamentals of AI and Data Science and their applications has become crucial not just for industry professionals but also for educators who are moulding the next generation of innovators and thinkers. AI-driven systems are now being used in almost all services we use in our everyday life, including social media, banking, e-commerce, navigation and conversational agents (e.g. ChatGPT). These helpful aids now have profound applications in diverse fields, including computer science, economics, development, business, education, engineering and law. As these technological innovations drive changes, it is more important than ever to learn (i) new pedagogical approaches for teaching AI and Data Science courses, (ii) ways of using generative AI as a thought companion and co-teacher, and (iii) grounding AI and Data Science concepts with real-world applications. This in-depth and hands-on 2-day workshop is specially designed for educators, aiming to unravel the core concepts in artificial intelligence and data science, teaching strategies that cater to the needs of the students as they prepare for the industry and academia, new pedagogical innovations, and to provide insights into their real-world applications.

Learning Outcomes	
	By the end of this workshop, the participants should be able to:
LO1:	Teach foundational concepts and principles of AI and Data Science and their relevance in today's technological landscape.
LO2:	Ground AI and Data Science teachings with real-world applications, ensuring students are well-prepared for both industry and

	academia.
LO3:	Engage with hands-on activities and strategies designed to cater to the specific needs of students, enhancing their learning experience.
LO4:	Build data pipelines (collection, cleaning, EDA, modelling, evaluation, results) for “repeatable” work.
LO5:	Use Python and popular data science libraries like Pandas, Numpy, Matplotlib, and Scikit-learn to analyse and visualise data.
LO6:	Utilise generative AI as a collaborative tool in the teaching process, harnessing its potential as a thought-companion and co-teacher.
LO7:	Communicate results effectively

Summary	
Key Concepts/Topics Covered	**Timestamp**
Day-1	
[A Pedagogical Framework for Teaching AI & Data Science] • How can we integrate computational and data-driven approaches to teach AI and data science effectively? [Understanding the Data Science Process] *What is data science? Why is it significant? What are the different pedagogical approaches for teaching data science?* • Exploring various shades of data • Leveraging data for effective communication • Viewing data science as creative problem-solving	Session I
[Descriptive Statistics, Data Acquisition, and Sampling Strategies]	Session II

What are some techniques for sound data design? How can we effectively summarise data? • Statistics of data and deceptive descriptions • Approaches to data design: random and non-random sampling strategies • Case Study [Data Cleaning, Exploratory Data Analysis, and Visualisation] *How can we transform data for ease of analysis? How do we systematically conduct data explorations?* • Principles of exploratory data analysis and visualisation • Analysing text data • Case Study	
[Experiments, Causal Inference and Foundations of Statistical Inference] *Is the relationship between X (e.g., smoking) and Y (e.g., cancer) causal? What is the uncertainty in my data?* • Experiments and Observational Studies • Impact Evaluations and Evidence-based Policy-making • Hypothesis Testing, Confidence Intervals, and p-values • Case Study	Session III
[Machine Learning] *Can I use X to predict Y? Will my predictions generalise to unseen scenarios?* • Simple and Multiple Linear Regression • Feature Engineering and Bias-Variance Tradeoff • Gradient Descent • Cross Validation and Regularization • Classification, Logistic Regression, and Evaluation Metrics ○ Common pitfalls and gaps in understanding around decision boundaries and evaluation metrics • Case Study [Big Data Analytics] *What is Big Data? What are Ethics? Why should we care about the Ethics of Data?*	Session III & IV

• Spark, TensorFlow, and PyTorch • Case Study	
Day-2	

[Classifiers: All for one… and one for all!]
• AI and Machine Learning
  ○ AI vs. Machine Learning
  ○ Machine Learning vs. Traditional Computer Science
• It's all about classifiers
  ○ From simple cat vs. dog image classification to ChatGPT
• Unsupervised Learning: Classification without labels
• From the basic to most advanced ML:
  ○ From Logistic Regression to Neural Networks,
  ○ From Neural Networks to Transformers/ChatGPT
	Session V

[Non-Parametric Classifiers]
• K Nearest Neighbors (KNN)
• Advancements of KNN and Decision Trees
• Silver Bullets: Ensemble methods
  ○ Decreasing Variance using Bagging
  ■ Random Forests
  ○ Decreasing Bias using Boosting
  ■ Gradient Boosted Trees
	Session VI

[Challenges: Fairness, Ethics, Responsible AI and Data Science] • Ethics and the Data Science Lifecycle • Bias and Fairness Criteria • Risks of AI ○ Inaccuracies	Session VII

○ Societal Biases ○ Impact on the Job Market	
[ChatGPT as a Co-teacher] ● The affordances of Conversational Agents ● ChatGPT as a thought-companion for preparing lectures ● ChatGPT as a writing aid while compiling notes, assignments and exams ● ChatGPT as a creative artist for designing lecture slides ● ChatGPT as an insightful grading assistant ● Challenges and Risks of unplanned integration	Session VIII

Reading Material
Will be shared at the workshop.