Professional Certificate in AI in Financial Crime Compliance

Compliance and Governance in AI for Financial Crime Detection

Actionable alerts
: Alerts generated by AI systems that require further investigation by compliance officers due to a high likelihood of financial crime.

Anomaly detection
: The use of AI to identify unusual patterns or outliers in data that may indicate financial crime.

Artificial intelligence (AI)
: The simulation of human intelligence processes by machines, especially computer systems. These processes include learning (the acquisition of information and rules for using the information), reasoning (using rules to reach approximate or definite conclusions), and self-correction.

Anti-Money Laundering (AML)
: A set of procedures, laws, and regulations designed to stop the practice of generating income through illegal actions.

Challenger banks
: Financial institutions that provide online-only banking services and are often more agile and innovative than traditional banks.

Compliance
: The process of ensuring that an organization adheres to laws, regulations, and guidelines relevant to its operations, including those related to financial crime.

Computer vision
: A field of AI that trains computers to interpret and understand the visual world.

Deep learning
: A subset of machine learning that uses artificial neural networks with many layers (i.e., deep) to learn and make decisions.

Explainability
: The ability to understand and interpret the decisions made by AI systems.

False negative
: An instance where the AI system fails to detect financial crime that is present.

False positive

: An instance where the AI system incorrectly detects financial crime.

Financial Crime
: Any illegal activity carried out by using or hiding the proceeds of illegal acts. Common examples include money laundering, fraud, and corruption.

Fraud detection
: The use of AI to identify and prevent fraudulent activities.

General Data Protection Regulation (GDPR)
: A regulation in EU law on data protection and privacy in the European Union and the European Economic Area.

Genetic algorithms
: A subset of evolutionary algorithms that use techniques inspired by natural selection, such as mutation, crossover, and selection.

Governance
: The establishment of policies, and continuous monitoring of their proper implementation, to ensure that compliance objectives are met.

Know Your Customer (KYC)
: The process of a business verifying the identity of its clients and assessing their suitability, along with the potential risks of illegal intentions towards the business relationship.

Machine learning
: A subset of AI that enables machines to learn and improve from experience without being explicitly programmed.

Natural language processing (NLP)
: A field of AI that focuses on the interaction between computers and human language, in particular how to program computers to process and analyze large amounts of natural language data.

Neural networks
: A type of machine learning model inspired by the human brain, composed of interconnected layers of nodes or "neurons".

Regtech
: The use of technology, particularly AI, to enhance regulatory processes.

Regulatory compliance
: The process of adhering to laws, regulations, and guidelines relevant to an organization's operations.

Risk scoring
: The process of assigning a score to individuals or entities based on their risk level of being involved in

financial crime.

Supervised learning
: A type of machine learning where the model is trained on a labeled dataset, allowing it to learn the relationship between inputs and outputs.

Synthetic identity fraud
: A type of fraud where a criminal creates a new identity using a combination of real and fake information.

Text mining
: The process of transforming unstructured text data into structured data for further analysis.

Transaction monitoring
: The process of analyzing financial transactions to detect and prevent financial crime.

Unsupervised learning
: A type of machine learning where the model is trained on an unlabeled dataset, allowing it to find patterns and relationships in the data without prior knowledge.

Unstructured data
: Data that doesn't have a predefined structure or format, such as text, images, and videos.

Feature engineering
: The process of creating new features or variables from existing data to improve the performance of machine learning models.

Fuzzy matching
: A technique used to identify strings that are approximately equal to a given pattern, often used in name matching to account for spelling variations and errors.

Latent Dirichlet allocation (LDA)
: A generative statistical model that allows sets of observations to be explained by unobserved groups that explain why some parts of the data are similar.

Libor (London Interbank Offered Rate)
: A benchmark interest rate that some of the world's leading banks charge each other for short-term loans.

Named entity recognition (NER)
: A subtask of NLP that seeks to locate and classify named entities in text into predefined categories such as person names, organizations, locations, medical codes, time expressions, quantities, monetary values, percentages, etc.

Natural language understanding (NLU)
: A subfield of NLP that focuses on the machine reading comprehension of text.

One-class SVM

: A variation of the Support Vector Machine algorithm used for anomaly detection, where the model is trained on normal data and used to detect abnormal data.

Outlier detection
: The identification of data points that are significantly different from other data points.

Principal component analysis (PCA)
: A statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components.

Random forest
: An ensemble learning method that operates by constructing multiple decision trees at training time and outputting the class that is the mode of the classes (classification) or mean/average prediction (regression) of the individual trees.

Rule-based systems
: AI systems that make decisions based on a set of predefined rules.

Support vector machines (SVM)
: A supervised machine learning algorithm that can be used for both classification or regression challenges. However, it is mostly used in classification problems.

Term Frequency-Inverse Document Frequency (TF-IDF)
: A numerical statistic used to reflect how important a word is to a document in a collection or corpus.

Word embeddings
: A technique in NLP where words or phrases from the vocabulary are mapped to vectors of real numbers.

YOLO (You Only Look Once)
: A real-time object detection system that is able to detect objects in images and videos.