

Key Findings/Summary

AWS Machine Learning Foundations Course

1. **Machine Learning Definition:** The scientific study of algorithms and statistical models to perform a task using inference instead of instructions.
 2. **Machine Learning Types:** Supervised, unsupervised, and reinforcement learning.
Machine Learning Challenges: Data quality, representativeness, sufficiency, overfitting, and underfitting.
 3. **Machine Learning Definition:** A scientific study of algorithms and statistical models that perform tasks through inference rather than explicit instructions.
Business Applications of Machine Learning: Spam detection, recommendation systems, and fraud detection.
 4. **Types of Machine Learning:** Categorised into three main types, though the specific types are not mentioned in the provided text.
 5. **Machine Learning Types:** Supervised, unsupervised, and reinforcement learning.
 6. Machine Learning Challenges: Data quality, business complexity, user expertise, and technology integration.
 7. **Machine Learning Implementation:** Defining objectives, data collection, model evaluation, and deployment using Amazon SageMaker.
 8. **Course Objective:** Provides foundational knowledge and practical skills in machine learning for internships and career advancement.
 9. **Data Collection and Security:** Covers data sources (private, communicable, open source), ETL with AWS Glue, and security measures like AWS IAM and CloudTrail.
 10. **Data Evaluation:** Includes data description using descriptive statistics and visualisation techniques like histograms.
 11. **Data Cleaning Techniques:** Handling missing data using `dropna()` and imputation methods (mean/median for symmetric/skewed data) and outlier detection using box plots and scatterplots.
-

-
- 12. Feature Selection Methods:** Filter (Pearson correlation, LDA, ANOVA, Chi-Square), Wrapper (forward/backward selection), and Embedded (decision trees, Lasso, Ridge) methods.
 - 13. Model Deployment Options:** Amazon SageMaker hosting services and Best Transformation for deploying trained and tuned models.
 - 14. Performance Metrics:** Sensitivity, specificity, accuracy, precision, F1 score, ROC, AUC-ROC, mean squared error, mean absolute error.
 - 15. Hyperparameter Tuning:** Model, optimiser, and data hyperparameters are used to define the model, learning process, and data attributes.
 - 16. Forecasting:** Predicting future values based on historical data, considering trends, seasonality, cyclic patterns, and irregularities.
 - 17. Time Series Handling:** Methods include forward/backward fill, moving average, interpolation, downsampling (mean/sum), and upsampling.
 - 18. Data Smoothing:** Used to remove outliers and anomalies, improving data quality for modelling and production.
 - 19. Time Series Algorithm Selection:** Factors like stationarity, trends, and autocorrelation influence the choice of algorithms like ARIMA, DEEP AR+, ETS, NPTS, and Prophet.
 - 20. Computer Vision Definition:** Automated extraction of information from digital images.
 - 21. Computer Vision Applications:** Public safety, autonomous driving, medical imaging, manufacturing process control, and more.
 - 22. Face Detection in Computer Vision:** Involves bounding boxes, face attributes, emotions, facial landmarks, quality, pose, confidence score, and comparison with a target.
 - 23. Natural Language Processing (NLP):** Computational algorithms for analysing and representing human language, addressing challenges like precision, context-based meaning, and lack of structure.
 - 24. Text Pre-processing:** Steps include removing stop words, normalising text, standardising unrecognised text, encoding, and spelling/grammar checks.
 - 25. Tokenisation and Feature Engineering:** Involves loading data using tokens, developing features with models like bag of words, term frequency, inverse document frequency, and TF-IDF, and capturing context through word tagging and entity extraction.
 - 26. Amazon Translate Use Cases:** International websites, software localisation, multilingual content, chatbots, and international media.
-

-
- 27. Amazon Comprehend Use Cases:** Document analysis, fraud detection, mobile app analysis, and content management.
 - 28. Amazon Lex Use Cases:** Inventory and sales management, interactive assistance, customer service interfaces, and database queries.
 - 29. Generative AI Definition:** Based on foundational models to create new content.
 - 30. Generative AI Use Cases:** Enhanced customer experience, boosted employee productivity, optimised processes, and enhanced creativity and content creation.
 - 31. AWS Services for Generative AI:** Code generation (Amazon Q developer), foundational model as a service (Amazon Bedrock, Amazon SageMaker Jumpstart), and computing (AWS Titanium, AWS Inferentia).

Certificate & Badge:

