## Data Science Training Topics Syllabus

| | |
|---|---|
| Introduction to Data Science | Understanding the data science process<br>Roles and responsibilities in data science |
| Data Collection and Data Sources | Data types and sources<br>Data collection methods<br>Data quality and cleaning |
| Data Exploration and Visualization | Exploratory data analysis (EDA)<br>Data visualization tools and techniques<br>Descriptive statistics |
| Data Preprocessing and Feature Engineering | Data cleaning and preprocessing<br>Handling missing data<br>Feature selection and engineering |
| Machine Learning Fundamentals | Supervised, unsupervised, and reinforcement learning<br>Model selection and evaluation<br>Overfitting and underfitting |
| Linear and Logistic Regression | Linear regression models<br>Logistic regression models<br>Model evaluation and interpretation |
| Classification and Clustering Algorithms | Decision trees and random forests<br>Support vector machines (SVM)<br>k-means clustering and hierarchical clustering |
| Natural Language Processing (NLP) | Text preprocessing<br>Sentiment analysis<br>Named entity recognition |
| Time Series Analysis | Time series data handling<br>Forecasting techniques<br>Seasonal decomposition |
| Neural Networks and Deep Learning | Introduction to artificial neural networks<br>Deep learning architectures (CNNs, RNNs, LSTM)<br>Transfer learning |
| Model Evaluation and Hyperparameter Tuning | Cross-validation<br>Hyperparameter optimization<br>Bias-variance trade-off |
| Feature Selection and Dimensionality Reduction | Principal Component Analysis (PCA)<br>Feature importance techniques |

| | |
|---|---|
| | L1 and L2 regularization |
| Big Data and Distributed Computing | Introduction to big data technologies (Hadoop, Spark)<br>Distributed data processing |
| Data Science Tools and Libraries | Popular data science libraries (e.g., Python, R, pandas, scikit-learn)<br>Integrated development environments (IDEs) |
| Data Ethics and Privacy | Ethical considerations in data science<br>Data privacy regulations and compliance |
| Data Visualization and Communication | Communicating insights effectively<br>Data storytelling and visualization best practices |
| Capstone Project | Applying data science skills to a real-world project<br>Presentation and documentation |
| Industry Applications and Case Studies | Data science in various industries (e.g., healthcare, finance, marketing)<br>Real-world case studies |
| Career Development and Job Search | Building a data science portfolio<br>Resume and interview preparation |