KKR GENAI INNOVATIONS

IT PROFESSIONALS NON CODERS

Weeks 1-2

Foundations of AI & Data in Business & IT

Topics:

☑ Overview of AI's applications in business and IT.
☑ Key roles in Business Analytics, Data Analysis, Quality Assurance, and IT Support.
☑ Data types and structures: structured, semi-structured, unstructured.
☑ Data storage and retrieval basics.
☑ Data quality fundamentals: accuracy, consistency, completeness.
☑ Using Excel and Power BI for data organization and visualization.
☑ Creating basic charts and dashboards in Tableau.
☑ Understanding non-coding data handling techniques.
☑ Organizing data for analysis.
☑ Introduction to using data insights for business decisions.

Tools:

☑ Microsoft Excel: Data organization, basic analysis.
☑ Power BI: Data visualization and dashboard creation.
☑ Tableau: Interactive dashboards and data exploration.

Practical Exercises:

☑ Structuring and analyzing data in Excel.
☑ Building a dashboard in Power BI or Tableau.

Resources:

☑ Harvard Online - Data Science for Business: Data Science for Business
☑ YouTube:
Google Cloud AI in Business
Power BI Tutorial for Beginners
☑ Tableau Free Training: Tableau Learning

Weeks 3-4

Python Fundamentals for Data Analysis

Topics:

☑ Python basics: variables, data types, operators.
☑ Data structures: lists, dictionaries, tuples.
☑ Control flow: if-else, loops.
☑ unctions and modules in Python.
☑ Introduction to Pandas and NumPy for data handling.
☑ Loading and handling data from files (CSV, Excel).
☑ Using Jupyter Notebook for data analysis.
☑ Google Colab for cloud-based coding.
☑ Data manipulation basics with Pandas.
☑ Simple data transformations and summarizations.

Tools:

☑ Jupyter Notebook: Interactive coding environment.
☑ Google Colab: Cloud-based Python platform.
☑ VS Code: Python development environment.

Practical Exercises:

☑ Writing basic Python scripts for data manipulation.
☑ Exploring data structures and control flow in Jupyter Notebook.

Resources:

☑ YouTube:
☑ Corey Schafer's Python Tutorials
Jupyter Notebook Basics by Tech with Tim
Real Python - Pandas Guide: Real Python
☑ Google Colab Tutorials: Google Colab

Weeks 5-6

Data Cleaning & Transformation Techniques

Topics:

☑ Data cleaning techniques for accuracy and completeness.
☑ Handling missing data (imputation, deletion).
☑ Removing duplicates and outliers.
☑ Data normalization and standardization.
☑ Converting data types (e.g., strings to numbers, dates).
☑ Encoding categorical variables (one-hot, label encoding).
☑ Text data cleaning with regular expressions.
☑ Filtering and sorting data in Pandas.
☑ Aggregating data with groupby functions.
☑ Introduction to OpenRefine for advanced cleaning.

Tools:

☑ Pandas: Data cleaning and manipulation.
☑ OpenRefine: Advanced cleaning for complex datasets.

Practical Exercises:

☑ Cleaning and transforming datasets using Pandas.
☑ Performing advanced data cleaning with OpenRefine.

Resources:

☑ Kaggle - Data Cleaning Course: Kaggle
☑ YouTube:
Data Cleaning with Pandas by Data School
OpenRefine Basics

Weeks 7-8

Data Visualization Basics

Topics:

☑ Data visualization principles and chart types.
☑ Creating bar charts, line charts, histograms in Matplotlib.
☑ Generating scatter plots and box plots.
☑ Advanced statistical visualizations with Seaborn.
☑ Choosing appropriate chart types for data.
☑ Customizing visualizations (labels, titles, legends).
☑ Working with subplots for multi-chart visualization.
☑ Visualizing data relationships and distributions.
☑ Basics of color theory in data visualization.
☑ Exporting visualizations to files (PNG, PDF).

Tools:

☑ Matplotlib: Fundamental plotting.
☑ Seaborn: Advanced visualizations.
☑ Plotly: Interactive charts.

Practical Exercises:

☑ Creating and customizing visualizations in Matplotlib and Seaborn.
☑ Building an interactive chart with Plotly.

Resources:

☑ YouTube:
StatQuest Data Visualization
Data Visualization with Matplotlib by Sentdex
☑ DataCamp - Data Visualization with Python: DataCamp

Weeks 9-10

Introduction to Cloud Computing for AI and Data

Topics:

☑ Basics of cloud computing: IaaS, PaaS, SaaS models.
☑ Benefits of cloud computing for AI and data science.
☑ Key cloud providers: AWS, Azure, Google Cloud Platform (GCP).
☑ Understanding virtual machines and containers.
☑ Introduction to serverless computing and its use in AI.
☑ Overview of compute services in the cloud (EC2, Azure Vms, GCP Compute Engine).
☑ Overview of storage services (S3, Azure Blob, GCP Storage).
☑ Basics of cloud networking and security.
☑ Cloud data storage solutions: databases, data lakes.
☑ Introduction to cloud-based development environments.

Tools:

☑ AWS, Azure, GCP Free Tiers: Cloud services exploration.
☑ Docker: Containerization.

Practical Exercises:

☑ Setting up a VM on AWS or Azure.
☑ Deploying a Docker container for AI models.

Resources:

☑ AWS Training - Cloud Practitioner Essentials: AWS Training
☑ YouTube:
Intro to Cloud Computing - Google Cloud
Microsoft Azure Fundamentals
☑ Docker Get Started Guide: Docker

Weeks 11-12

AI/ML Services on Azure, AWS, & GCP

Topics:

☑ Introduction to Azure Cognitive Services (text, vision, speech).
☑ AWS SageMaker for model training and deployment.
☑ GCP Vertex AI overview for AI/ML workflows.
☑ Text analytics and language processing with Azure.
☑ Image recognition with AWS Rekognition.
☑ AutoML tools in cloud platforms for automated model building.
☑ Comparing AI/ML service offerings across cloud platforms.
☑ Overview of deployment options (model endpoints, batch predictions).
☑ Data storage and retrieval for ML on each platform.
☑ Using cloud-based notebooks (SageMaker, Azure Notebooks).

Tools:

☑ Azure Cognitive Services: Text, vision, and language processing.
☑ AWS SageMaker: ML lifecycle management.
☑ GCP Vertex AI: Model building and deployment.

Practical Exercises:

☑ Deploying an AI model on AWS SageMaker.
☑ Using Azure's Text Analytics API for sentiment analysis.

Resources:

☑ AWS Machine Learning University: AWS ML University
☑ YouTube:
Azure Cognitive Services Overview
GCP's Vertex AI Tutorial

Weeks 13-14

Data Science Workflow & Advanced Data Wrangling

Topics:

☑ Overview of the data science workflow.
☑ Data gathering techniques and sources (APIs, scraping).
☑ Data cleaning in a structured workflow.
☑ Exploratory data analysis and visualization for insights.
☑ Advanced data merging and joining techniques.
☑ Working with timestamps and datetime objects.
☑ Filtering and slicing data in Pandas.
☑ Data transformation with pivot and melt functions.
☑ Exporting clean data for model building.
☑ Documenting data workflows for collaboration.

Tools:

☑ Pandas: Advanced data manipulation.
☑ BeautifulSoup/Requests: Web scraping.
☑ SQL: Querying databases.

Practical Exercises:

☑ Combining data from multiple sources with Pandas.
☑ Writing SQL queries to clean and organize data.

Resources:

☑ DataCamp - Data Manipulation with Pandas: DataCamp
☑ YouTube:
Advanced Pandas Techniques by Keith Galli
Corey Schafer - Web Scraping

Weeks 15-16

Introduction to Large Language Models (LLMs) & Prompt Engineering

Topics:

☑ Fundamentals of Transformer models (GPT, BERT).
☑ Understanding tokenization and embeddings in LLMs.
☑ Applications of LLMs in NLP tasks (text generation, Q&A, summarization).
☑ Introduction to APIs for LLMs (OpenAI, Hugging Face).
☑ Basics of prompt engineering: crafting effective prompts.
☑ Strategies for prompt optimization (clarity, brevity).
☑ Introduction to fine-tuning LLMs for specific tasks.
☑ Understanding context and coherence in LLM responses.
☑ Limitations and ethical considerations of LLMs.
☑ Using LLMs for summarization, translation, and Q&A.

Tools:

☑ OpenAI API: Accessing GPT models.
☑ Hugging Face Transformers: Pre-trained models for NLP.
☑ GPT-3 Playground: Experimenting with prompts.

Practical Exercises:

☑ Crafting prompts to perform NLP tasks.
☑ Using Hugging Face to fine-tune a language model.

Resources:

☑ OpenAI API Documentation: OpenAI
☑ YouTube:
Two Minute Papers - GPT-3 Overview
Hugging Face Transformers Course
☑ Hugging Face Transformers Course: Hugging Face

Weeks 17-18

Exploratory Data Analysis (EDA) & Feature Engineering

Topics:

☑ Importance of exploratory data analysis in data science.
☑ Analyzing data distributions and patterns.
☑ Detecting and handling outliers.
☑ Creating custom features (feature engineering).
☑ Feature scaling techniques (normalization, standardization).
☑ Encoding categorical variables.
☑ Feature selection methods.
☑ Dimensionality reduction techniques (PCA, t-SNE).
☑ EDA with visualizations for insights.
☑ Reporting EDA findings effectively.

Tools:

☑ Pandas Profiling: Automated EDA reports.
☑ Scikit-Learn: Feature engineering tools.
☑ Plotly: Visualizing feature distributions.

Practical Exercises:

☑ Performing EDA on a dataset.
☑ Feature engineering on a dataset.

Resources:

☑ Kaggle - Exploratory Data Analysis: Kaggle
☑ YouTube:
StatQuest - EDA
Data School - Feature Engineering

Weeks 19-20

Supervised Learning Techniques

Topics:

☑ Introduction to supervised learning: labeled data and predictions.
☑ Linear regression and logistic regression.
☑ Support Vector Machines (SVM).
☑ Decision Trees and Random Forests.
☑ Model evaluation metrics (accuracy, precision, recall, F1).
☑ Cross-validation techniques and train-test split.
☑ Avoiding overfitting and underfitting.
☑ Regularization techniques (L1, L2).
☑ Hyperparameter tuning with grid search.
☑ Model interpretability basics.

Tools:

☑ Scikit-Learn: Supervised ML algorithms.
☑ MLFlow: Experiment tracking.
☑ SHAP: Model interpretability.

Practical Exercises:

☑ Training and evaluating a supervised model.
☑ Experiment tracking with MLFlow.

Resources:

☑ Coursera - Supervised ML by Andrew Ng: Coursera
☑ YouTube:
StatQuest - Supervised Learning
☑ Scikit-Learn Documentation: Scikit-Learn

Weeks 21-22

Unsupervised Learning & Dimensionality Reduction

Topics:

☑ Introduction to unsupervised learning (no labeled data).
☑ Clustering techniques (K-means, hierarchical).
☑ Association rule learning for pattern recognition.
☑ Dimensionality reduction methods (PCA, t-SNE).
☑ Visualization of high-dimensional data.
☑ Evaluating clustering quality (silhouette score).
☑ Feature extraction for complex datasets.
☑Anomaly detection with clustering.
☑ Comparing clustering algorithms.
☑ Applications of unsupervised learning in business.

Tools:

☑ Scikit-Learn: Unsupervised learning and dimensionality reduction.
☑ TSNEViewer: Visualizing t-SNE results.
☑ ELKI: Advanced clustering and outlier detection.

Practical Exercises:

☑ Applying K-means clustering on a dataset.
☑ Reducing dimensions with PCA and visualizing results.

Resources:

☑ Fast.ai - Practical ML Course: Fast.ai
☑ YouTube:
StatQuest - Clustering and PCA
3Blue1Brown - PCA Visual Guide

Weeks 23-24

Model Deployment & Evaluation in the Cloud

Topics:

☑ Basics of model deployment for production.
☑ Building REST APIs for ML models.
☑ Introduction to Flask and FastAPI for model deployment.
☑ Monitoring model performance over time.
☑ Setting up automated model retraining.
☑ Model performance metrics in production (latency, accuracy).
☑ Security best practices for deployed models.
☑ Introduction to Prometheus for monitoring.
☑ Using Grafana for visualization.
☑ Scaling models in production with cloud solutions.

Tools:

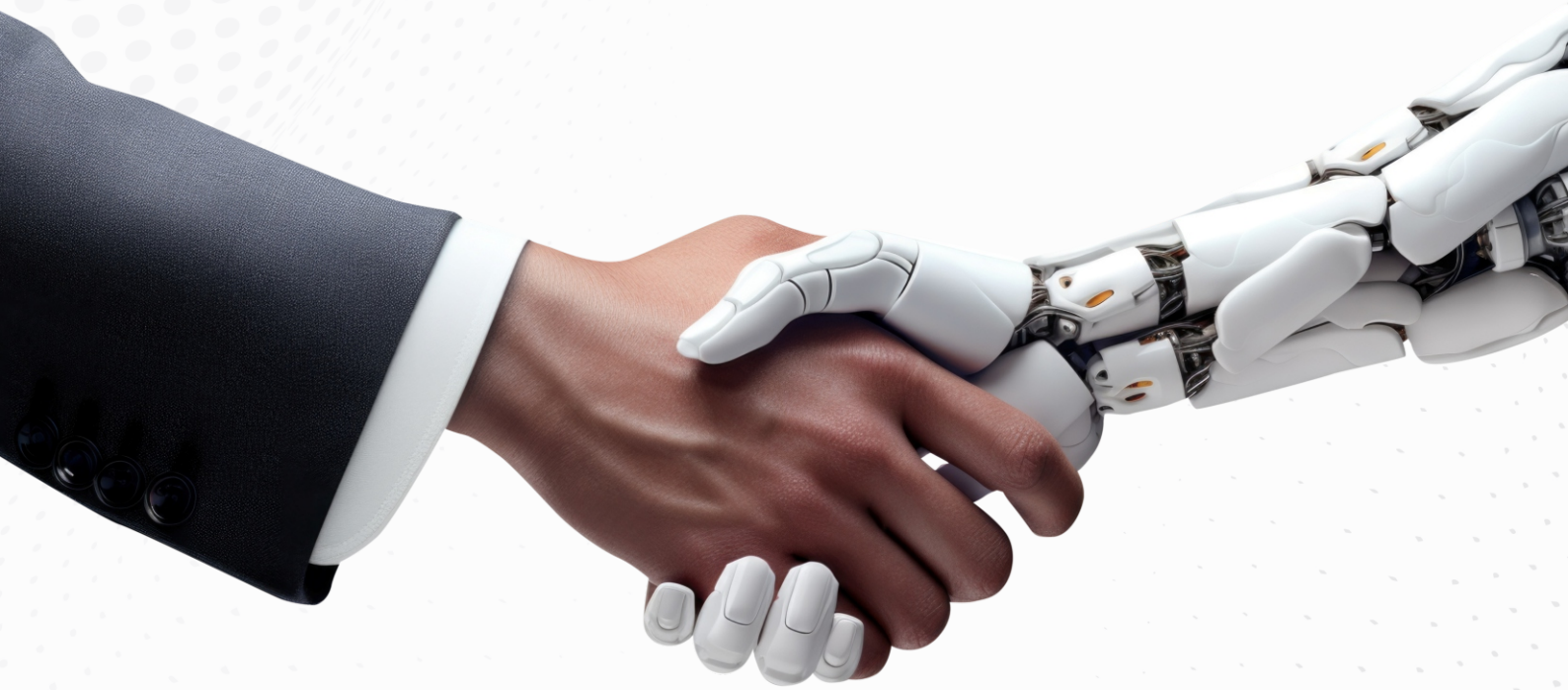
☑ Flask/FastAPI: Deploying APIs for models.
☑ Prometheus and Grafana: Monitoring tools.
☑ Docker: Containerizing models for deployment.

Practical Exercises:

☑ Deploying a model with Flask/FastAPI as a REST API.
☑ Setting up Prometheus and Grafana to monitor a deployed model.

Resources:

☑ AWS SageMaker Model Deployment: AWS
☑ YouTube:
Flask API for ML Model Deployment
Prometheus and Grafana Monitoring



https://www.kkrgenaiinnovations.com/enrol

Follow Us on Social Media

+1 4708616312 / +91 89770 81409

Info@kkrgenaiinnovations.com

www.kkrgenaiinnovations.com

11535 Park Woods Circle Suite A3
Alphareta GA 30005