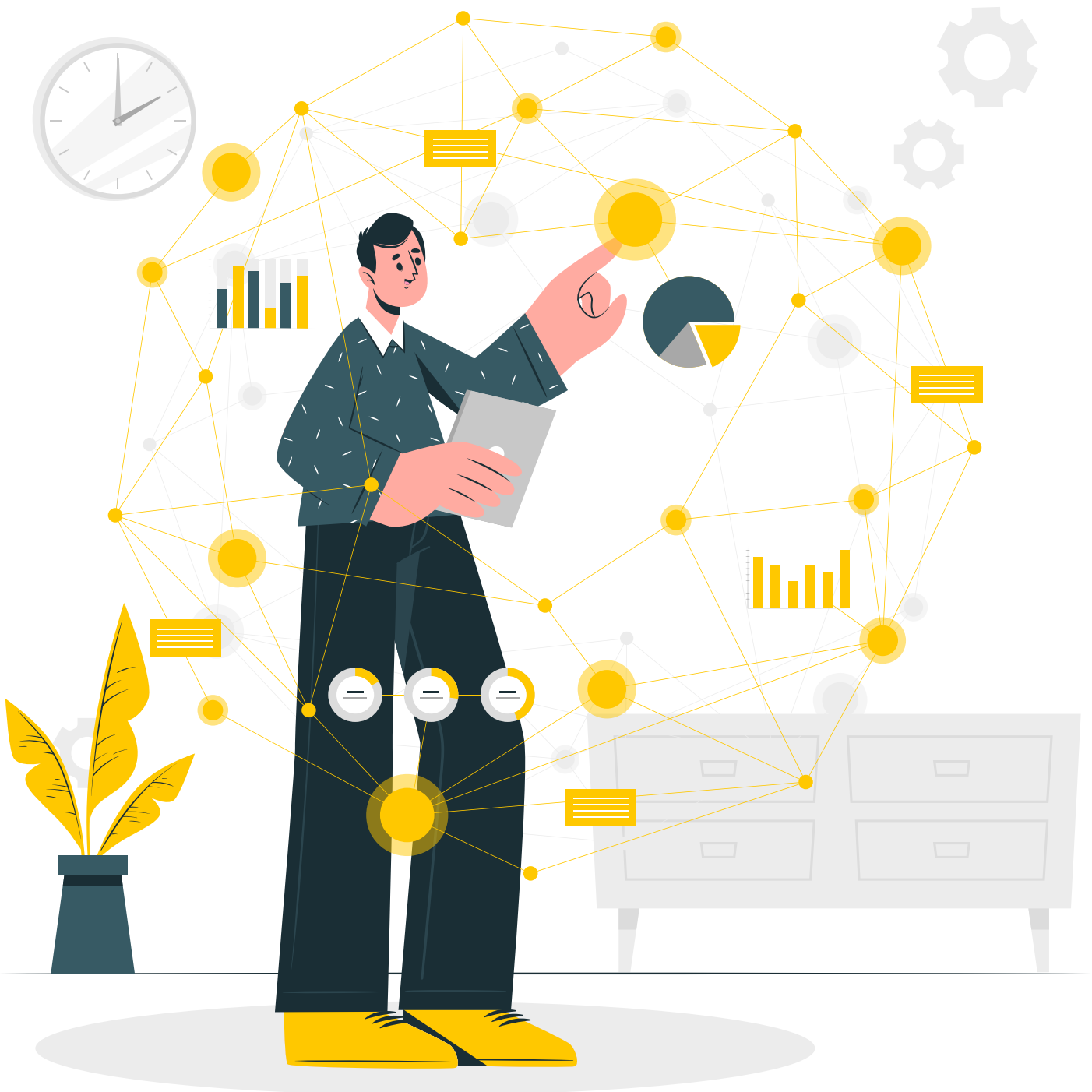# Data Science with AI Masters Program

Jobpakado

Build your career with us

JobPakado is a custom path for students and learners to achieve a 100% successful career in their dream field. With a vision and professional team, we crafted job-ready skills programs with industry-relevant skills to make it possible for every student to get their dream job.

Instructor-Led training

Facilitated learning with live instructor guiding participants skill development.

Career Oriented Sessions

Professional growth workshops, skills training, career counseling.

Project & Exercises

We will conduct exercises everythime after finishing the every module

24*7 Support

Round-the-clock doubt clarification support every day in the week.

Executive Alumni Status

We provide best in class mentorship by industry experts.

Assured Interviews

Guaranteed industrial interviews, optimized career. Best results assured.

Profile Building

We will help to build best profile to stand unique from the others.

Mock Interview Preparation

We'll conduct free mock interviews to build your self confidence.

Placement Services

We provide best Placements to the best learners in top Companies.

Self-paced videos

We'll provide Self-paced videos for reference

1:1 Career Mentoring Sessions

We provide 1 to 1 mentoring to give the best career.

Certification

We provide certificate from our Organization

More than 500 students have turned their dream career into reality. You can be the next with the India's top 100% job assistance programs

You're ready for this. Trust in yourself and your abilities. We believe in you!

Why to choose this data science domain ?

- Huge Vacancies
- Overwhelming demand
- High pay package
- Stable employment
- Simple to grasp
- Rapid Growth

Top Hiring Companies

Google   amazon   IBM   HSBC

Microsoft   Infosys   accenture   Deloitte.

Top Job Roles

- Data Scientist
- Machine learning Engineer
- Tableau Developer
- Business Analyst
- Data Analyst
- Data Engineer
- Business Intelligence Specialist
- Business Intelligence Professional

Tools & Technologies

python   NumPy   Pandas

matplotlib   seaborn   Power BI

tableau   MACHINE LEARNING   AI Artificial Intelligence

Office 365   ChatGPT   MySQL

Jobpakado

Our Program Index

Our Master's program is more focused on practicals than theory. It includes 5-6 months of online or offline training, and you'll work with experienced instructors who have 15 years of industry expertise. They will mentor you and help you get hired.

Introduction

- Introduction to Data Science
- Life cycle of Data Science
- Skills required for Data Science
- Applications of Data Science in different Industries

Python Programming

- Introduction to Python programming
- Installing Python and setting up the development environment
- Variables, data types, and type conversions
- Operators: arithmetic, comparison, logical, and assignment
- Control structures: if, else, elif, nested conditionals
- Loops: for loop, while loop, nested loops
- Functions and their role in code organization
- Handling errors with try-except blocks

Working with files: reading, writing, and appending data

Numpy

- Introduction to NumPy
- What is NumPy?
- Advantages of using NumPy for data science
- Installing NumPy
- NumPy Arrays
- Creating NumPy arrays
- Data types in NumPy arrays
- Array indexing and slicing
- Array operations (arithmetic, broadcasting, etc.)

Jobpakado

Working with Multi-dimensional Arrays

- Creating multi-dimensional arrays
- Shape, size, and dimensions of arrays
- Array reshaping and resizin

Array Operations

- Element-wise operations
- Aggregation and reduction functions
- Array broadcasting and vectorization

Array Manipulation

- Adding and removing elements in arrays
- Stacking and concatenating arrays
- Splitting arrays

Indexing and Slicing

- Boolean indexing and masking
- Fancy indexing and advanced slicing

Array Math and Statistics

- Mathematical functions in NumPy
- Statistical functions (mean, median, variance, etc.)
- Linear algebra with NumPy (dot product, matrix operations)

Random Number Generation

- Generating random data with NumPy
- Seeds and reproducibility

Handling Missing Data with NumPy

- Dealing with NaN values in NumPy arrays

File Input and Output

- Loading and saving NumPy arrays from/to files (CSV, text, binary)

Pandas

Introduction to Pandas

- What is Pandas and why it is essential for Data Science

- Installing Pandas

- Importing the Pandas library

- Data structures in Pandas: Series and DataFrame

Getting Started with DataFrames

- Creating DataFrames from different data sources (CSV, Excel, etc.)

- Inspecting the DataFrame: head(), tail(), info(), describe()

- Selecting and accessing data: loc, iloc

- Conditional data selection using boolean indexing

Data Cleaning and Preparation

- Handling missing data: isnull(), dropna(), fillna()

- Removing duplicates: drop_duplicates()

- Data transformation: applying functions, mapping, and replacing values

- Data type conversions

Data Manipulation with Pandas

- Indexing and slicing DataFrames

- Filtering rows and columns based on conditions

- Aggregation and grouping data: groupby()

- Pivot tables and cross-tabulations

- Merging, joining, and concatenating DataFrames

Introduction to Data Visualization

Understanding the importance of data visualization in data science
Overview of popular Python visualization libraries: Matplotlib and Seaborn

Matplotlib Basics

- Installing Matplotlib and getting started
- Creating basic plots: Line plot, Scatter plot, Bar plot
- Customizing plots: Labels, Titles, Colors, Markers, and Linestyles
- Working with multiple plots: Subplots

Advanced Matplotlib

- Plotting categorical data: Bar plots, Pie charts
- Plotting time series data
- Working with different plot types: Histograms, Box plots, Violin plots
- Adding annotations and text to plots

Seaborn Introduction

- Installing Seaborn and understanding its advantages over Matplotlib
- Seaborn vs. Matplotlib: When to use each library

Seaborn Plots

- Creating basic Seaborn plots: Scatter plot, Line plot, Bar plot
- Visualizing distributions: Histograms, KDE plots
- Plotting categorical data: Bar plots, Box plots, Violin plots
- Heatmaps and cluster maps for correlation analysis

Data Exploration and Visualization

- Combining Matplotlib and Seaborn for more complex visualizations
- Exploring relationships between variables using joint plots and pair plots
- Using FacetGrids for multi-plot visualizations

Styling and Aesthetics

- Customizing Seaborn styles and themes
- Setting plot aesthetics: Colors, Fonts, Grids
- Incorporating external color palettes

Jobpakado

Introduction to Power BI

Introduction to Power BI

- What is Power BI?
- Key features and benefits
- Understanding the Power BI ecosystem

Getting Started with Power BI

- Downloading and installing Power BI Desktop
- Creating a Power BI account
- Overview of the Power BI interface

Importing and Transforming Data

- Connecting to data sources (e.g., Excel, databases, online services)
- Data loading and transformation using Power Query Editor

Data Modeling in Power BI

- Understanding tables and relationships
- Creating calculated columns and measures

Introduction to DAX (Data Analysis Expressions) language

- Creating Visualizations
- Building basic visualizations (e.g., bar charts, pie charts, tables)
- Applying formatting and customization to visuals
- Utilizing interactive features (e.g., drill-down, cross-filtering)

Designing Reports and Dashboards

- Combining multiple visuals to create reports
- Organizing and arranging visuals on a report page
- Creating interactive dashboards for data exploration

Statistics and Probability for Data Science

Introduction to Data Science and Descriptive Statistics
- Overview of data science and its applications
- Types of data: categorical, numerical, continuous, and discrete
- Data visualization: histograms, box plots, scatter plots, etc.
- Measures of central tendency: mean, median, mode
- Measures of dispersion: variance, standard deviation, range

Probability Basics
- Fundamentals of probability theory
- Probability rules: addition and multiplication rules
- Conditional probability and Bayes' theorem
- Probability distributions: discrete and continuous
- Expectation and variance of random variables

Sampling and Estimation
- Simple random sampling
- Point estimation and interval estimation
- Sampling distributions
- Central Limit Theorem and its applications

Statistical Inference
- Hypothesis testing: null and alternative hypotheses
- Type I and Type II errors
- Commonly used significance levels
- One-sample and two-sample t-tests
- Chi-square test for independence
- Analysis of Variance (ANOVA)

Regression Analysis
- Linear regression
- Multiple regression
- Model evaluation and selection
- Residual analysis

Probability Distributions for Data Science

- Binomial distribution

- Poisson distribution

- Normal distribution and its applications

- Exponential distribution

Bayesian Statistics

- Bayesian inference

- Bayesian updating

- Bayesian regression

Machine Learning

Introduction to Data Science and Machine Learning

- Understanding the data science workflow

- Overview of machine learning and its applications

- Types of machine learning: supervised, unsupervised, and reinforcement learning

Data Preprocessing

- Data cleaning and handling missing values.

- Feature engineering and selection

- Data transformation and normalization

- **Supervised Learning Algorithms**

- Linear regression

- Logistic regression

- Decision trees and random forests

- Support Vector Machines (SVM)

- k-Nearest Neighbors (k-NN)

- Naive Bayes

- Gradient Boosting algorithms (e.g., XGBoost, LightGBM)

Unsupervised Learning Algorithms

- K-means clustering

- Hierarchical clustering

- Dimensionality reduction techniques (e.g., Principal Component Analysis - PCA)

Model Evaluation and Validation

- Cross-validation techniques

- Performance metrics: accuracy, precision, recall, F1-score, etc.

- Overfitting and underfitting

Model Selection and Hyperparameter Tuning

- Grid search and randomized search for hyperparameter optimization

Ensemble Methods

- Bagging and bootstrap aggregating

- Stacking and blending

Introduction to Recommender Systems

- Collaborative filtering

- Content-based filtering

Introduction to Time Series Analysis

- Time series data handling

- Time series forecasting using machine learning

Introduction to AI

- Basics of neural networks: Neurons, activation functions, and backpropagation
- Introduction to TensorFlow and Keras libraries
- Building and training deep learning models for image classification and
- regression tasks
- Convolutional Neural Networks (CNN) for image analysis
- Recurrent Neural Networks (RNN) for sequential data processing

Natural Language Processing (NLP)
- Introduction to NLP and its applications
- Text preprocessing: Tokenization, stop word removal, stemming, and lemmatization
- Bag-of-Words (BoW) and TF-IDF representation
- Sentiment analysis and text classification using NLP techniques

Advance Excel using MS365 with ChatGPT

Data Cleaning and Preparation:
- Removing duplicates
- Handling missing values
- Data validation and data integrity
- Text-to-columns and data formatting
- Data transformation techniques

- **Advanced Formulas and Functions**
- Logical functions (IF, AND, OR, etc.)
- Lookup and reference functions (VLOOKUP, HLOOKUP, INDEX-MATCH)
- Text functions (CONCATENATE, LEFT, RIGHT, MID, etc.)
- Date and time functions
- Statistical functions (AVERAGEIF, COUNTIF, SUMIF, etc.)

Jobpakado

Data Analysis Tools
- PivotTables and PivotCharts
- Data filtering and sorting
- Subtotals and grouping data
- What-If analysis (Data tables, Goal Seek, Scenario Manager)

Data Visualization
- Creating charts and graphs (line charts, bar charts, pie charts, etc.)
- Customizing chart elements
- Advanced charting techniques (combination charts, sparklines, etc.)

Data Analysis with Excel
- Descriptive statistics (mean, median, standard deviation, etc.)
- Histograms and frequency distributions
- Correlation and regression analysis
- Data analysis using Excel's Data Analysis ToolPak add-in

Data Import and Export
- Importing data from external sources (CSV, text files, databases, etc.)
- Exporting data to different file formats (CSV, PDF, etc.)

Automating Tasks with Macros
- Introduction to VBA (Visual Basic for Applications)
- Recording and running macros
- Creating custom functions and procedures

Data Integration with Excel
- Combining data from multiple sources
- Data consolidation techniques

Excel and Data Visualization Tools Integration:
- Linking Excel data with Power BI or Tableau
- Importing and exporting data between Excel and Python/R

Excel Efficiency and Productivity:
- Keyboard shortcuts and time-saving tips
- Excel add-ins and productivity tools

Jobpakado

MYSQL

Introduction to SQL

- Overview of SQL and its applications in data science.
- Relational databases and their components.
- Basic SQL syntax and commands.

Querying Data

- SELECT statement and retrieving data from tables.
- Filtering data with the WHERE clause.
- Sorting data with the ORDER BY clause.
- Limiting results with the LIMIT clause.

Filtering and Combining Data

- Using logical operators (AND, OR, NOT) to create complex queries.
- Working with NULL values.
- Combining data from multiple tables with JOIN operations (INNER JOIN, LEFT JOIN, RIGHT JOIN, FULL JOIN).

Data Aggregation

- Understanding aggregate functions (SUM, AVG, COUNT, MIN, MAX).
- GROUP BY clause and grouping data based on specific criteria.
- HAVING clause for filtering aggregated data.

Data Manipulation

- Inserting, updating, and deleting data in a database.
- Transactions and their importance.

Data Analysis with SQL

- Subqueries and their applications.
- Common table expressions (CTEs).
- Window functions for advanced data analysis.

Jobpakado

Advanced SQL Techniques

- Working with date and time data.
- String manipulation and regular expressions.
- Conditional expressions and CASE statements.

Performance Optimization

- Understanding indexes and their impact on query performance.
- Query optimization techniques.

Data Import and Export

- Importing data into a SQL database.
- Exporting query results to external files.

Working with Large Datasets

- Partitioning data to improve performance.
- Using temporary tables and table variables.

Real Time Projects

Sentiment Analysis for Social Media Comments

Summary: Implement a deep learning model to perform sentiment analysis on social media comments. The model will classify comments as positive, negative, or neutral, enabling businesses to gauge customer sentiments effectively. **Domain: Natural Language Processing (NLP).**

Predicting Customer Churn in Telecom Industry

Summary: Leveraging machine learning algorithms, this project predicts customer churn in the telecom industry. By analysing customer behaviour and usage patterns, it helps identify potential churners, allowing businesses to take proactive measures to retain valuable customers. **Domain: Customer Analytics, Telecom.**

Stock Price Prediction using Time Series Analysis

Summary: Employing time series forecasting techniques and historical stock data, this project predicts future stock prices. It can aid investors and traders in making informed decisions and managing their portfolios effectively. **Domain: Finance, Stock Market.**

Image Classification for Medical Diagnostics

Summary: Using deep learning and convolutional neural networks (CNNs), this project aims to classify medical images into different diagnostic categories. It can assist medical professionals in making accurate and timely diagnoses, leading to improved patient outcomes. **Domain: Medical Imaging, Healthcare.**

Personalized Movie Recommendation System

Summary: This project builds a recommendation system using collaborative filtering and matrix factorization techniques to suggest movies based on users' preferences and viewing history. It enhances user experience on streaming platforms and boosts user engagement. **Domain: Recommender Systems, Entertainment.**

Jobpakado

Crop Yield Prediction for Precision Agriculture

Summary: Leveraging machine learning and historical agricultural data, this project predicts crop yields for different regions and seasons. It aids farmers in optimizing resource allocation and improving overall crop productivity. **Domain: Agriculture, Remote Sensing.**

Recommender System for E-commerce

Summary: Implement a personalized recommender system that suggests products to users based on their past behaviour and preferences. This project aims to enhance user engagement and increase sales in e-commerce platforms. **Domain: Recommendation Systems.**

Air Quality Prediction

Summary: Develop a machine learning model to predict air quality levels in each area based on historical data and environmental factors. This project aims to provide early warnings for potential air pollution events.
Domain: Environmental Sciences.

Natural Language Generation for Chatbots

Summary: Implement a natural language generation model to improve the conversational abilities of chatbots. The system will generate human-like responses, enhancing user interactions.
Domain: Natural Language Processing (NLP) and Chatbots.

Language Translation using Sequence-to-Sequence Models

Summary: This project involves implementing a sequence-to-sequence model for language translation tasks. By leveraging recurrent neural networks or transformers, the model can translate text from one language to another, facilitating cross-lingual communication in the **natural language processing domain.**

Climate Change Impact Analysis using Machine Learning

Summary: Build a machine learning model to analyse climate data and predict the potential impact of climate change on specific regions. The project contributes to climate research, environmental planning, and disaster management. **Domain: Environmental Science, Climate Change Analysis**

Are you ready to succeed in **your career?**

Apply now

www.jobpakado.com

Jobpakado