SKILL INTERN

DATA SCIENCE

Designed for skill proficiency

PROGRAM HIGHLIGHITS

- **Basic-Advanced Level Training**
  By Experienced Mentors
- **Live & Recorded Lectures**
  At Your Flexible Schedule
- **Real Time Projects**
  Minor & Major Projects
- **Accredited certificates**
  Program approved ISO Certificate
- **Internships**
  Opportunities will be provided
- **Placement Guidance**
  Assistance from industrial EXPERTS

OUR MOTIVE

UPSKILL

Empowering Minds For Tomorrow

ENHANCE

Discover Your Next Ambition

MOTIVATE

Empowering Minds, Igniting Futures

ABOUT US

Skill Intern is a leading EdTech company dedicated to empowering engineering students with the skills and knowledge necessary to excel in today's competitive job market. Our mission is to bridge the gap between theoretical learning and practical application, enabling students to develop a strong foundation and enhance their employability.

Data Science has emerged as one of the most critical fields in the modern digital age due to its profound impact on various aspects of business, technology, and everyday life. Here are several key reasons why Data Science is important

WHY DS ?

- Data-Driven Decision Making
- Competitive Advantage
- Operational Efficiency
- Personalization
- Risk Management
- Innovation and Research
- Improving Quality of Life
- Enhancing Business Strategies
- Handling Big Data
- Societal Impact
- Job Opportunities
- Enhanced Decision-Making Capabilities

LEARNING PATH

- Introduction to Data Science
- Python for Data Science
- Data Handling and Manipulation
- Data Visualization
- Statistics and Probability
- Machine Learning Fundamentals
- Advanced Machine Learning
- Time Series Analysis
- Deep Learning
- Natural Language Processing (NLP)
- Big Data and Cloud Computing
- Data Science Project Management

Module 1: Introduction to Data Science

- Definition and Overview of Data Science
- The Data Science Lifecycle
- Roles and Responsibilities of a Data Scientist
- Applications of Data Science in Various Industries

Module 2: Python for Data Science

- Introduction to Python Programming
- Python Basics: Variables, Data Types, Operators
- Control Structures: Conditionals and Loops
- Functions and Modules
- Introduction to Jupyter Notebooks

Module 3: Data Handling and Manipulation

- Importing Data: CSV, Excel, JSON, SQL
- Data Manipulation with Pandas
- Data Cleaning: Handling Missing Values, Duplicates
- Data Transformation: Aggregation, Grouping, Merging
- Exploratory Data Analysis (EDA)

Module 4: Data Handling and Preprocessing

- Types of Data: Structured and Unstructured
- Data Collection and Cleaning
- Exploratory Data Analysis (EDA)

Feature Engineering and Selection

Module 5: Machine Learning Basics

- Introduction to Machine Learning
- Supervised vs. Unsupervised Learning
- Classification Algorithms (K-Nearest Neighbors, Decision Trees)
- Regression Algorithms (Linear Regression, Polynomial Regression)

Module 6: Advanced Machine Learning

- Ensemble Methods (Random Forest, Gradient Boosting)
- Clustering Algorithms (K-Means, Hierarchical Clustering)
- Dimensionality Reduction (PCA, LDA)
- Model Evaluation and Validation (Cross-Validation, ROC Curve)

Module 7: Neural Networks and Deep Learning

- Introduction to Neural Networks
- Perceptrons and Multi-Layer Perceptrons
- Backpropagation and Gradient Descent
- Introduction to Deep Learning Frameworks (TensorFlow, Keras)

Module 8: Deep Learning Architectures

- Convolutional Neural Networks (CNNs)
- Recurrent Neural Networks (RNNs)
- Long Short-Term Memory Networks (LSTMs)
- Autoencoders and Generative Adversarial Networks (GANs)

Module 9: Natural Language Processing (NLP)

- Introduction to NLP
- Text Preprocessing (Tokenization, Stemming, Lemmatization)
- Bag of Words and TF-IDF
- Advanced NLP Techniques (Word Embeddings, Transformers)

Module 10: Natural Language Processing (NLP)

- Introduction to NLP
- Text Preprocessing: Tokenization, Stemming, Lemmatization
- Bag of Words and TF-IDF
- Advanced NLP Techniques: Word Embeddings, Transformers
- Sentiment Analysis and Text Classification

Module 11: Big Data and Cloud Computing

- Introduction to Big Data Technologies
- Hadoop Ecosystem: HDFS, MapReduce, Hive
- Introduction to Spark and PySpark
- Cloud Platforms for Data Science: AWS, Azure, Google Cloud
- Data Storage and Management in the Cloud

Module 12: Data Science Project Management

- Project Planning and Execution
- Data Science Workflow and Best Practices
- Version Control with Git and GitHub
- Collaboration Tools and Techniques
- Presentation and Communication of Results

Assignments & Assessments

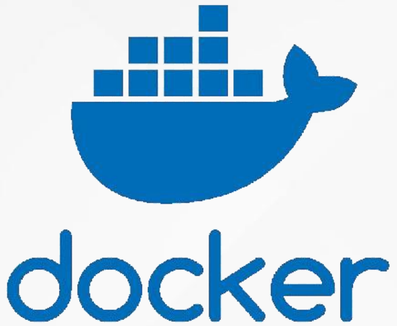
- Weekly assignments based on module topics
- Mid-term project: Wireframing and prototyping a small application
- Final project: Comprehensive DATA SCIENCE project
- Participation in class discussions and activities

Recommended Reading

- "Python for Data Analysis" by Wes McKinney
- "Data Science from Scratch: First Principles with Python" by Joel Grus
- "An Introduction to Statistical Learning: with Applications in R" by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani

FRAME WORKS

TOOLS USED



DOCKER

JUPTER NOTEBOOK

***In case of additional tools used, It will be discussed in live class**

CERTIFICATIONS

www..skillintern.com

THANK YOU

www.skillintern.com