Machine Learning Boot Camp / SkillJourney / Part 1: Data Prep & Cleaning (TTMI5510)

Category: AI, Data & Analytics | **Vendor:** Trivera Tech

Duration: 24.00 hours (3 days) **19.5 CPD Hours** **Rating:** ⭐ 4.6 (5,878 reviews)

Course Information

Delivery Format: Instructor Led - Online

Course Overview

Our®Machine Learning Essentials Boot Camp:®Preparing Your Data®is a three-day hands-on skills immersion course geared for students who need to how to effectively prepare and optimize data for use in machine learning models, ensuring they produce accurate, useful and insightful predictions. ®

About This Course

In the world of machine learning, the quality of input data is critical. Machine learning models that use bad data input produce inaccurate and unreliable results, undermining their effectiveness and trustworthiness. Our Machine Learning Essentials Boot Camp: Preparing Your Data is a three-day hands-on skills immersion course geared for students who need to how to effectively prepare and optimize data for use in machine learning models, ensuring they produce accurate, useful and insightful predictions.

Throughout the course, guided by our expert instructor, you'll engage in workshop-style practical labs that will provide you with the real-world skills and hands-on experience needed to manage, prep and clean your data for successful machine learning model applications.

You'll learn how to translate diverse data into an analytically-friendly format, ensuring compatibility with machine learning algorithms. You'll learn how to scale and normalize data, ensuring consistent data representation, crucial for accurate model training and predictions. You'll navigate the intricacies of data transformation and refinement, and learn how to translate diverse datasets into formats friendly to machine learning algorithms. You'll also explore feature selection and dimensionality reduction, striking the balance between data richness and computational efficiency. You'll also grasp how to safeguard your data's journey with robust pipelines and preventive measures against data leakage, cementing the trustworthiness of your real-world model deployments. Lastly, you'll explore the complete lifecycle of a machine learning project, from data preparation to model deployment, you're equipped to oversee and implement comprehensive data-driven solutions.

By the end of this immersive boot camp, you'll be fully-equipped with a comprehensive skillset that not only enhances the predictive power of your models but also sets the foundation for innovative, data-driven solutions. You'll be ready to advance in your Machine Learning journey, leveraging your newly acquired skills towards model proficiency.

Who Should Attend

This course is geared for data scientists and business professionals seeking to leverage data insights in decision-making. It's also ideal for software developers wanting to diversify their skills into the exciting field of machine learning. Whether you're a student eager to jumpstart your career or an experienced professional looking to enhance your data-driven strategies, our hands-on workshop offers a valuable learning experience to transform you into a confident data handler and problem-solver.

Prerequisites & Entry Requirements

General Prerequisites:

This is an intermediate-level program, designed to prepare attendees for a deeper dive into next-level, heavy hands-on machine learning courses and workshops. Attendees should have practical, hands-on experience working with Python for Data Science, pandas and numpy.

Learning Outcomes

Upon successful completion of this course, participants will be able to:

Throughout the course you will explore:
-Data Encoding: Dive into data encoding to seamlessly translate diverse information into a machine-friendly format.
-Data Manipulation Mastery: You'll get comfortable with encoding, scaling, and normalizing data. By the end of the course, the curse of dimensionality will no longer be a challenge.
-Quality Analysis Confidence: Learn how to identify and remove duplicates, handle null values, manage outliers, and work with dates in your data. You'll be a pro at maintaining clean datasets.
-Feature Analysis Wizardry: Discover how to identify unused columns, detect low variance ones, and understand multicollinearity. By the end of the workshop, feature selection will feel like second nature.
-Pipeline Proficiency: Gain a deep understanding of the critical role of pipelines in machine learning and develop the skills to create and implement your own data preprocessing pipelines.
-Machine Learning Basics: Get introduced to the fundamentals of machine learning, understand k-fold cross-validation, master the art of partitioning data, and learn how to prevent data leakage. You'll be set to step confidently into the world of machine learning.

Detailed Course Outline

Getting Started with Data

Explore the role and importance of data in machine learning.

Encoding data: Transform raw data into a format suitable for analytics.

Dealing with the curse of dimensionality: Navigate high-dimensional spaces effectively.

Scaling and normalizing data: Standardize data for consistent analysis.

Hands-on Activity / Lab

Structural Analysis

Delve into the intricate patterns that define data.

Importing libraries: Equip yourself with the right tools for data manipulation.

Importing data: Initiate the first steps of data-driven exploration.

Conducting basic data investigation: Peek into the essence of your dataset.

Utilizing relevant tools for data structure analysis: Get acquainted with state-of-the-art tools to dissect data structure.

Hands-on Activity / Lab

Quality Analysis

Refine data sets by spotting and fixing errors.

Identifying and removing duplicates: Ensure uniqueness in your dataset.

Handling null values and missing data: Fill the gaps in your data with precision.

Detecting and managing outliers: Understand and manage extreme data points.

Working with dates in data: Harness the power of time-series data.

Hands-on Activity / Lab

Exploratory Data Analysis

Dive deep into data to extract meaningful insights.

Conducting univariate analysis: Analyze one variable at a time.

Conducting bivariate analysis: Discover relationships between two variables.

Conducting multivariate analysis: Understand complex data interactions.

Using pivot tables for data analysis: Summarize data visually and numerically.

Understanding correlation: Measure linear relationships between variables.

Understanding mutual information: Gauge dependency between variables.

Hands-on Activity / Lab

Data Features

Pinpoint the most impactful data components.

Identifying and dropping unused columns: Streamline data for efficiency.

Detecting and handling low variance or no variance columns: Maintain data variability.

Understanding multicollinearity (VIF): Ensure independent predictor variables.

Feature Selection

Prioritize the most relevant data features for robust models.

Using wrappers (RFE, Forward, Backward selection): Implement dynamic feature selection.

Using filters (Statistical tests): Opt for features based on statistical relevance.

Using embedded methods: Integrate feature selection into algorithm functionality.

Understanding unsupervised feature selection methods: Navigate feature selection without target variables.

Hands-on Activity / Lab

Feature Importance

Gauge the significance of different data features in prediction.

Understanding dimensionality reduction: Simplify data without losing information.

Using Principal Component Analysis (PCA): Transform data to highlight variance.

Using Linear Discriminant Analysis (LDA): Optimize class separability.

Hands-on Activity / Lab

Encoding, Scaling, and Skewness

Tailor data formats for better compatibility with machine learning algorithms.

Encoding categorical variables: Convert categories into numerical values.

Scaling numerical variables: Maintain consistency in data magnitude.

Detecting and correcting skewness in data: Normalize data distributions.

Hands-on Activity / Lab

Pipelines

Streamline machine learning workflows with seamless data transitions.

Understanding the role of pipelines in machine learning: Appreciate the significance of efficient workflows.

Creating and implementing data preprocessing pipelines: Process data in a structured manner.

Using pipelines for efficient cross-validation and hyperparameter tuning: Optimize model parameters with ease.

Hands-on Activity / Lab

Introduction to Machine Learning

Lay the groundwork for next-level machine learning practices.

Understanding k-fold cross-validation: Assess model performance effectively.

Using resampling techniques: Balance dataset disparities.

Dividing data into training and test sets: Create a structured environment for model training and evaluation.

Identifying and preventing data leakage: Maintain the integrity of your datasets.

Understanding the basic types and applications of machine learning models

Capstone Project: Develop an end-to-end machine learning model: Apply the course skills to develop a complete data-driven projects.

Additional Course Details

Nexus Humans Machine Learning Boot Camp / SkillJourney / Part 1: Data Prep & Cleaning (TTMI5510) training program is a workshop that presents an invigorating mix of sessions, lessons, and masterclasses meticulously crafted to propel your learning expedition forward.

This immersive bootcamp-style experience boasts interactive lectures, hands-on labs, and collaborative hackathons, all strategically designed to fortify fundamental concepts.

Guided by seasoned coaches, each session offers priceless insights and practical skills crucial for honing your expertise. Whether you're stepping into the realm of professional skills or a seasoned professional, this comprehensive course ensures you're equipped with the knowledge and prowess necessary for success.

While we feel this is the best course for the Machine Learning Boot Camp / SkillJourney / Part 1: Data Prep & Cleaning (TTMI5510) course and one of our Top 10 we encourage you to read the course outline to make sure it is the right content for you.

Additionally, private sessions, closed classes or dedicated events are available both live online and at our training centres in Dublin and London, as well as at your offices anywhere in the UK, Ireland or across EMEA.

Frequently Asked Questions

Q: What delivery options are available for Machine Learning Boot Camp / SkillJourney / Part 1: Data Prep & Cleaning (TTMl5510)?

We offer multiple delivery formats:

- Live Instructor-Led Classroom Online (Virtual/Live Online)
- Traditional Instructor-Led Classroom Training (ILT)
- On-site delivery at your offices anywhere in United Kingdom
- Private dedicated courses customized for your team

Q: How many CPD hours does this course provide?

The 3-day Machine Learning Boot Camp / SkillJourney / Part 1: Data Prep & Cleaning (TTMl5510) course provides up to 19.5 CPD hours of structured learning. CPD certificates can be provided upon request.

Q: What is the duration of the Machine Learning Boot Camp / SkillJourney / Part 1: Data Prep & Cleaning (TTMl5510) training?

The training takes place over 3 day(s), with each day lasting approximately 24.00 hours including breaks for lunch and refreshments.

Q: Do you provide corporate training for Machine Learning Boot Camp / SkillJourney / Part 1: Data Prep & Cleaning (TTMl5510)?

Yes, we provide corporate training, dedicated training, and closed classes for Machine Learning Boot Camp / SkillJourney / Part 1: Data Prep & Cleaning (TTMl5510). Training can take place anywhere in United Kingdom including London, Manchester, Birmingham, Edinburgh, or live online allowing teams from across United Kingdom or internationally to attend.

Q: What related terms do people search for?

Popular related searches include: AI; Machine Learning

Q: Why choose Nexus Human for Machine Learning Boot Camp / SkillJourney / Part 1: Data Prep & Cleaning (TTMl5510)?

Nexus Human is recognized as one of the leading training providers. Our trainers have won multiple awards including:

- Small Firms Best Trainer Award
- National Training Partner of the Year (Ireland) - Multiple Years
- Global Top 30 Instructor Awards (2012, 2019, 2021)
- Tech Excellence Award Nominations
- Learning Performance Institute (LPI) External Training Provider Sponsor 2024

Q: Are there any discount codes available?

Yes! Use discount code **PENPAL5** when booking your Machine Learning Boot Camp / SkillJourney / Part 1: Data Prep & Cleaning (TTMl5510) training. Please note that only one discount code can be used per booking and cannot be combined with other special offers.

Nexus Human

Professional Training & Development

Email: info@nexushuman.com

Website: www.nexushuman.com

Phone: +353 1 889 8700 (Ireland) | +44 20 8142 8961 (UK)

© 2026 Nexus Human. All rights reserved. This brochure was generated on 13/11/2026.