

Regression vs Classification in Machine Learning Explained!

Home > Career

> Regression vs Classification in Machine Learning Explained!



[avcontentteam](#)

27 Jun, 2023 • 7 min read

As [data scientists](#) and experienced technologists, professionals often seek clarification when tackling machine learning problems and striving to overcome data discrepancies. It is crucial for them to learn the correct strategy to identify or develop models for solving equations involving distinct variables. Thus, understanding the disparity between two fundamental algorithms, Regression vs Classification, becomes essential. Classification and regression are both techniques employed in [machine learning](#), but they serve different purposes and are suited for distinct types of problems. Let's explore the difference between classification and regression in machine learning.

[Table of contents](#)

What is Regression?

Regression algorithms predict continuous value from the provided input. A supervised learning algorithm uses real values to predict quantitative data like income, height, weight, scores or probability. Machine learning engineers and data scientists mostly use regression algorithms to operate distinct labeled datasets while mapping estimations.

What is Classification?

A procedure in which a model or a function separates the data into discrete values, i.e., multiple classes of datasets using independent features, is called classification. A form If-Then rule derives the mapping function. The values classify or forecast the different values like spam or not

Regression vs Classification in Machine Learning Explained!

discrete label includes predicting the possibility of an actor visiting the mall for a promotion, depending on the history of the events. The labels will be Yes or No.



Source: [Analytics Vidhya Youtube Channel](#)

Types of Regression

1. Linear Regression

Most preferable and simple to use, it applies linear equations to the datasets. Using a straight line, the relationship between two quantitative variables i.e., one independent and another dependent, is modeled in simple [linear regression](#). A dependent variable's multiple linear regression values can use more than two independent variables. It is applicable to predict marketing analytics, sales, and demand forecasting.

2. Polynomial Regression

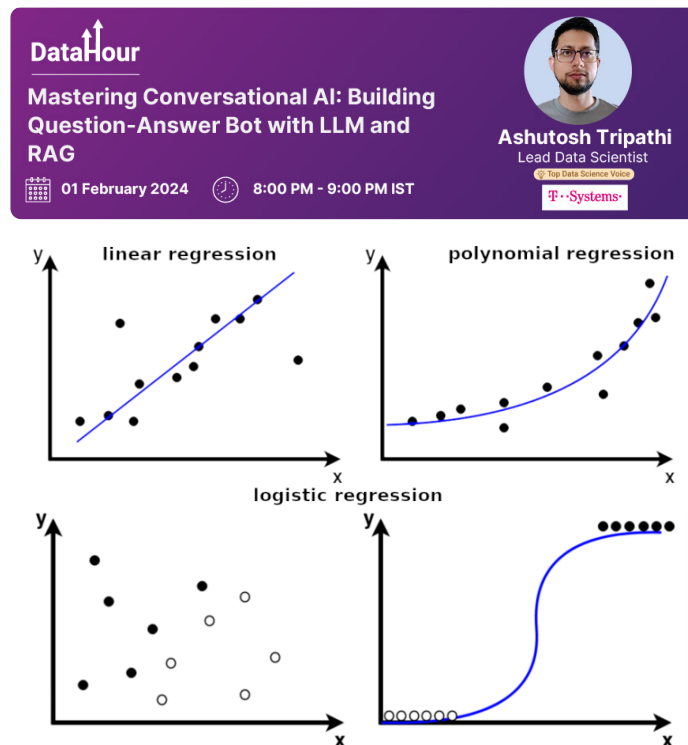
To find or model the [non-linear relationship](#) between an independent and a dependent variable is called polynomial regression. It is specifically used for curvy trend datasets. Various fields like social science, economics, biology, engineering and physics use a polynomial function to predict the model's accuracy and complexity. In ML, polynomial regression is applicable to predict customers' lifetime values, stock and house prices.

3. Logistic Regression

Commonly known as the logit model, [Logistic Regression](#) understands the probable chances of the occurrence of an event. It uses a dataset comprising independent

Regression vs Classification in Machine Learning Explained!

Classification.



Source: KapernikovAn

Types of Classification

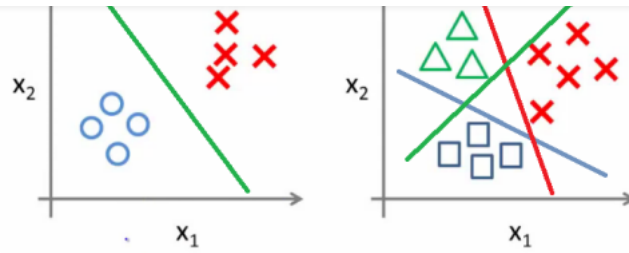
1. Binary Classification

When an input provides a dataset of distinct features describing each point, the output of the model delivered will be [binary](#) labeled representing the two classes i.e., categorical. For example, Yes or No, Positive or Negative.

2. Multi-class Classification

In machine learning, multi-class classification provides more than two outcomes of the model. Their subtypes are one vs all/rest and multi-class [classification algorithms](#). Multiclass does not rely on binary models and classifies the datasets into multiple classes. At the same time, OAA/OAR represents the highest probability and score from separate binary models trained for each class.

Regression vs Classification in Machine Learning Explained!



Source: Cloud2data

3. Decision Trees

Decisions and their consequences are in a tree-based model, where nodes of the [decision tree](#) confirm each node and edges show the consequence of that particular



[Values in Data Analysis](#)

Applications of Regression

1. Predicting Stock Prices

Regression algorithms create mathematical relationships between the [stock price](#) and related factors to predict accurate model values using historical data, screening trends and patterns.

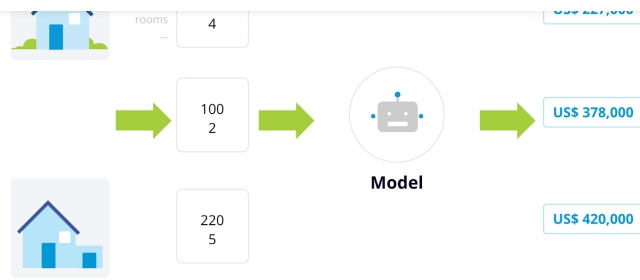
2. Sales Forecasting

Organizations planning sales strategies, inventory levels and marketing campaigns can use historical sales data, trends, and patterns to predict future sales. It helps [forecast sales](#) in wholesale, retail, e-commerce and other sales and marketing industries.

3. Real Estate Valuation

Establish mathematical equations to predict models that discover the values of real estate properties. An organization can easily determine the property price by depending on the amenities, size and location of the property along with its historical data, including market values and sale patterns. It is widely used by [real estate](#) professionals, sellers and buyers to assess expenses and investments.

Regression vs Classification in Machine Learning Explained!



Source: Tryolabs

Applications of Classification

1. Email Spam Filtering

Training is provided to the classifier using labeled data to classify the emails. Filtering of emails can be done by analyzing the two categorical data i.e., spam or not spam. The [filtered emails](#) are then automatically delivered to the appropriate class as per the selected features determined in the input.

2. Credit Scoring

Credit scores can be assessed using a classification algorithm. It analyses the history of the client, amount of transactions, loan sanctioned, income, demographic information and other factors to predict the informed decisions of loan approval for the applicants.

3. Image Recognition

Classifier is trained based on labeled data, thus assisting in predicting the images per the corresponding labeled class. The images with new content, like animals or objects, can easily be categorized into classes automatically.

Advantages and Disadvantages of Regression

Advantages

- **Valuable Insights:** Helps to analyze the relationships between distinct variables and achieve a significant understanding of the data.

Regression vs Classification in Machine Learning Explained!

values with high accuracy using independent variables.

- **Flexibility:** Regression is a flexible algorithm used to find or predict models of a wide range, and it includes logistic, linear, polynomial and many more.
- **Ease in Interpretation:** The analyzed results of regression can be visualized easily in the form of charts and graphical representations.

Disadvantages

- **False Assumptions:** The regression algorithm lies on numerous assumptions, thus resulting in false assumptions in the context of the real world. It includes normality of errors, linearity and independence.
- **Overfitting:** Inadequate performance can be applied to new and unseen data because the regression models are overly customized for the training data.
- **Outliers:** [Regression models](#) are sensitive to exceptions, thus, can have a significant effect on analyzed prediction results.

Advantages and Disadvantages of Classification

Advantages

- **Accuracy in Prediction:** With fitting training, the classification algorithm achieves high accuracy in the model prediction.
- **Flexible:** Classification algorithms have many applications like spam filtering, speech and image recognition.
- **Scalable Datasets:** Easy to apply in real-time applications that can scale up huge datasets easily.
- **Efficient and Interpretable:** The classification algorithm efficiently handles huge datasets and can classify them quickly, which is easy to interpret. It provides a better understanding of variable-to-outcome relationships.

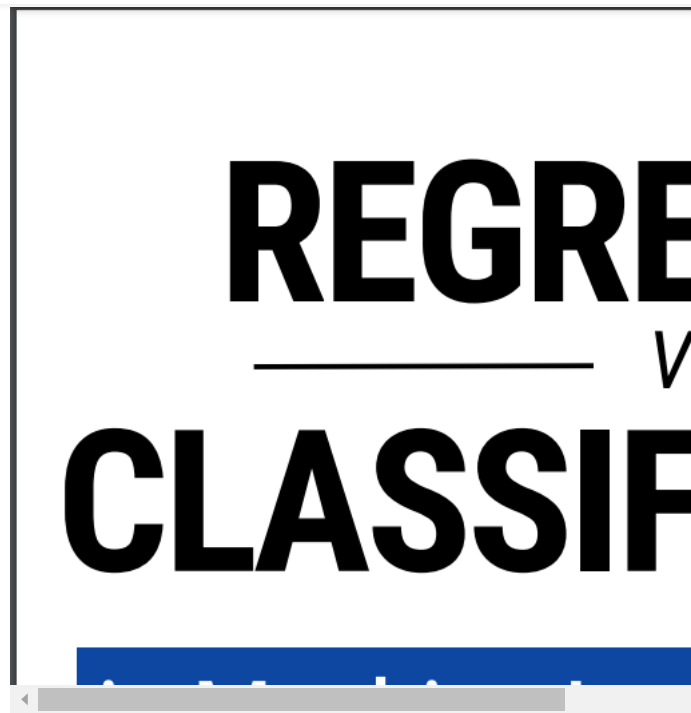
Regression vs Classification in Machine Learning Explained!

- **Bias:** If the training data does not represent the complete dataset, the classification algorithm may get biased with certain trained data.
- **Imbalanced Data:** If the classes of the datasets are not determined equally, the classification algorithm will read the majority and leave the minority class. For example, if there are two classes of data i.e., 85% and 15%, the majority class data will be represented as 85%, leaving the rest undefined.
- **Selection of Features:** Features must be defined in the classification algorithms, else the prediction of data is challenging with multiple or undefined features.

Differences Between Regression and Classification

Let us have a comparative analysis of regression vs classification:

Features	Regression	Classification
Main goal	Predicts continuous values like salary and age.	Predicts discrete values like stock and forecasts.
Input and output variables	Input: Either categorical or continuous Output: Only continuous	Input: Either categorical or continuous Output: Only categorical
Types of algorithm	Linear regression Polynomial regression Lasso regression Ridge regression	Decision trees Random forests Logistic regression Neural networks Support vector machines
Evaluation metric	R2 score Mean squared error Absolute error percentage error (MAPE)	Receiver operating characteristic curve Recall Accuracy Precision F1 score



[Complete Guide on Regression vs Classification](#) [Download Now](#)

When to Use Regression or Classification?

The classification vs regression usage in different domains is stated as follows:

A. Data types

[Data Types](#) used as input are continuous or categorical in regression and classification algorithms. But the target value in regression is continuous, whereas categorical is in the classification algorithm.

B. Objectives

Regression aims to provide accurate continuous values like age, temperature, altitude, stock prices, house rate, etc. The classification algorithm predicts class categories like a mail is either spam or not spam; the answer is either true or false.

C. Accuracy requirements

Regression mainly focuses on achieving the highest accuracy by decreasing the prediction errors like mean absolute error or mean squared error. On the other hand, classification focuses on achieving the highest accuracy

Regression vs Classification – End Note

Understanding the differences between regression vs classification algorithms is crucial for data scientists to solve market issues effectively. Accurate data predictions rely heavily on selecting the right models, ensuring high precision in the results. If you want to enhance your machine learning skills and become a true expert in the field, consider joining our [Blackbelt program](#). This advanced program offers comprehensive training and hands-on experience to take your data science career to new heights. With a focus on regression, classification, and other advanced topics, you'll gain a deep understanding of these algorithms and how to apply them effectively. Join the [program](#) today!

Frequently Asked Questions

Q1. What is the difference between classification and regression?

A. Classification and regression are machine learning tasks, but they differ in output. Classification predicts discrete labels or categories, while regression predicts continuous numerical values.

Q2. What is the difference between classification and regression loss?

A. Classification loss measures the error between predicted class probabilities and the true class labels, typically using cross-entropy loss. Regression loss, on the other hand, quantifies the difference between predicted continuous values and the actual values, often using mean squared error or mean absolute error.

Q3. What is the difference between regression and classification in predictive analysis?

A. In predictive analysis, regression focuses on predicting numerical outcomes, such as a house's price. On the other hand, classification aims to assign instances to predefined classes, like determining whether an email is spam. They serve different purposes based on the nature of the problem.

Regression vs Classification in Machine Learning Explained!

A. Regression predicts continuous numerical values, aiming to find relationships between variables. Classification assigns instances to discrete classes based on predefined criteria. Clustering is an unsupervised learning technique that groups similar models based on their features without predefined classes or continuous values.

Artificial Intelligence

classification

data science

deep learning

machine learning

regression



[avcontentteam](#)

27 Jun 2023

Career

Classification

Machine Learning

Regression

Responses From Readers

What are your thoughts?...

Submit reply

Related Articles



Analytics Vidhya's Top 10 Machine Learning Blogs in 2022

The article covers top 10 blogs on machine learning



Regression Analysis | Beginners Comprehensive Guide (Updated 2024)

Linear & logistic regression are different. Learn