MACHINE LEARNING ZOOMCAMP: 2024 COHORT

# **Brief Introduction to** [**DataTalks**](https://datatalks.club/)

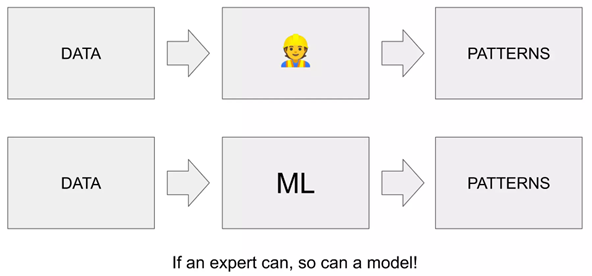
# **Introduction to Machine Learning**

## Introduction to Machine Learning

Overview:

1. Train a model
2. Use a model

This is a summary of what I’ve learned from the great ML course (<https://github.com/DataTalksClub/machine-learning-zoomcamp>) by Alexey Grigorev. All images from this post are from the course material. Images in other posts can also be copies of that material.

The introduction starts with an explanation about what ML really is. You can imagine a task that is normally done by an expert, like getting a good price for selling a car. The expert takes the data about the car and combines all the characteristics to get his opinion about the fair price. What he does is, he extracts patterns from the data.If a human is able to do this, so a model can do the same.

Following is an explanation of what ML is.

**Machine Learning** (ML) is about using ***features***and the **target**information to train a **model**and use this model to predict unknown object targets. In other words it is a process of extracting patterns from data (features + target).

To understand this, you have to distinguish between the terms feature, target and model.

**Features**means what we know about an object. In this example what we know about the characteristics of a car. A feature is a characteristic of an object in form of a number, string, or other more complex form (e.g. location information, …)

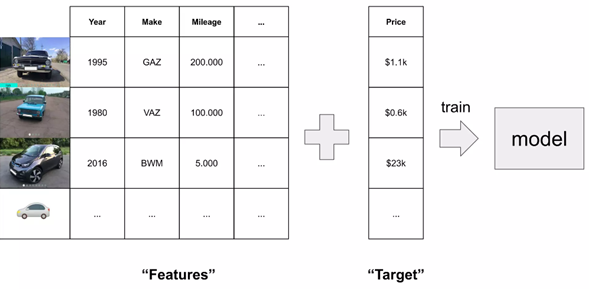
**Target**is what we want to predict. Other courses / sources also use the term label for this purpose.That means, in training, you talk about a labeled data set because you know the target. In this example, many labeled data sets of cars with prices are used to predict a label for an unknown data set of another car.

A **model**is the output artefact of a **training**that contains all the patterns learned from the trained examples. This output can be used later to make a **prediction**(try to output the target variable) based on features of an unknown object and the model itself.

**Train a Model**

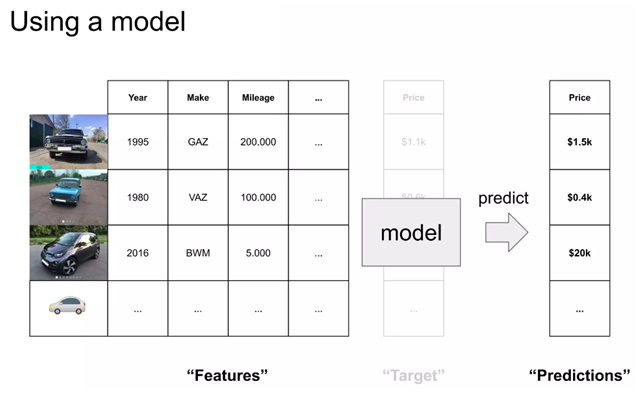
Model training is the process where the machine extracts the patterns from the given training data. In easy words the features are combined with the target – this leads to the model.

**Machine Learning**



**Use a Model**

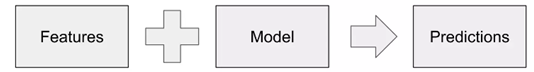
Mere training does not make a model useful. Only the application brings the benefit. Applying the trained model to an unknown data set (without target), you obtain a prediction for the missing information (here: the price).



To summarize the difference between model training



And prediction



is that in training process you use features and the target to get the model. And in the prediction process you only use the features and apply the trained model to get a prediction for the target variable.

## 1.2 ML vs Rule-Based Systems

## 1.3 Supervised Machine Learning

## 1.4 CRISP-DM

## 1.5 Model Selection Process

## 1.6 Setting up the Environment

## 1.7 Introduction to NumPy

## 1.8 Linear Algebra Refresher

## 1.9 Introduction to Pandas

## 1.10 Summary

## 1.11 Homework

# **Machine Learning for Regression**

# **Machine Learning for Classification**

# **Evaluation Metrics for Classification**

# **Deploying Machine Learning Models**

# **Decision Trees and Ensemble Learning**

# **Neural Networks and Deep Learning**

# **Serverless Deep Learning**

# **Kubernetes and TensorFlow Serving**