Python

For Machine Learning Applications

Ph.D. HOSNI Adil Imad Eddine

2022/2023

- Welcome to the presentation on Python for Machine Learning.
- In this presentation, we will explore the importance of Python in Machine Learning and how it can be used to build powerful and scalable Machine Learning models.
- We will cover the basics of Python, key Python libraries for Machine Learning, and some examples of Python-based Machine Learning projects.
- Whether you're new to Machine Learning or an experienced practitioner, this presentation will provide you with valuable insights into using Python for Machine Learning.
- All materials will be found in (here)

Table of Contents

What is machine learning

- ML is a branch of artificial intelligence. ML is programming computers to optimize a performance criterion using example data or past experience.
- This will Extract patterns, fit data to functions, classify data, etc
- ML systems can learn and improve with historical data, time and experience
- Learning is used when:
  - Human expertise does not exist (navigating on Mars),
  - Humans are unable to explain their expertise (speech recognition)
  - Solution changes in time (routing on a computer network)
  - Solution needs to be adapted to particular cases (user biometrics)

Applications

1 Machine learning

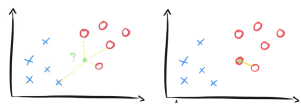
1. **Association** : $P(Y|X)$ probability that somebody who buys X also buys Y where X and Y are products/services.
2. **Supervised Learning**: For every example in the data there is always a predefined outcome. Models the relations between a set of descriptive features and a target
   - **Classification**: **Aka Pattern recognition**
   - **Regression**
3. **Unsupervised Learning**
4. **Reinforcement Learning**

Classification

1 Machine learning

Predicts which class a given sample of data (sample of descriptive features) is part of (discrete value). examples:

- Face recognition.
- Character recognition.
- Medical diagnosis.

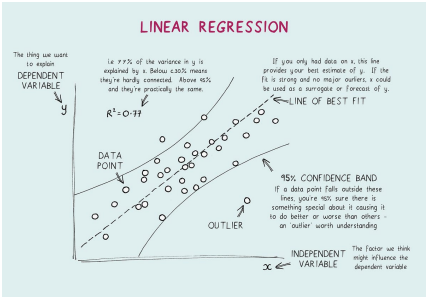


Regression

1 Machine learning

Linear Regression is a statistical method used to analyze the relationship between two or more variables by fitting a straight line to the observed data. Linear Regression is widely used in various fields such as:

- Navigating a car: Angle of the steering wheel.
- Price of a used car.
- Response surface design



Unsupervised Learning

There are not predefined and known set of outcomes. Look for hidden patterns and relations in the data.

A typical example: Clustering

Clustering

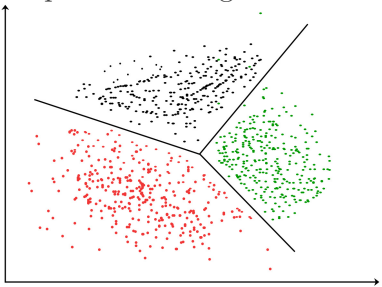


Table of Contents

What is Python ?

2 Why Python for machine learning?

Python is a high-level, general-purpose programming language that is easy to learn and use. It has a clear and concise syntax, and it emphasizes code readability and simplicity. Python is an interpreted language, meaning that code can be executed immediately without the need for compilation. It is widely used in various fields, including web development, data science, and machine learning, and it has a large and active community of developers and users.

Why Python for machine learning?

- Python is a popular and powerful programming language that is widely used in Machine Learning.
- Python is easy to learn and use, and it has a large and active community of developers and users.
- Python provides a wide range of libraries and tools that are specifically designed for Machine Learning, including NumPy, Pandas, Matplotlib, Scikit-learn, and TensorFlow.

Advantages of using Python

2 Why Python for machine learning?

Python is a popular choice for machine learning due to a number of factors. Here are some of the advantages of using Python for machine learning:

- Easy to Learn and Use: Python has a simple and intuitive syntax that is easy to learn, even for beginners. It also has a large and active community of developers and users who contribute to its ecosystem of tools and resources.
- Extensive Libraries and Frameworks: Python has a wide range of libraries and frameworks specifically designed for machine learning, such as NumPy, Pandas, Matplotlib, and Scikit-learn.
- Flexibility and Versatility: it a versatile language that can be applied to various fields, from web development to scientific computing.
- Open-Source.
- Large Community and Support.

Training path

2 Why Python for machine learning?

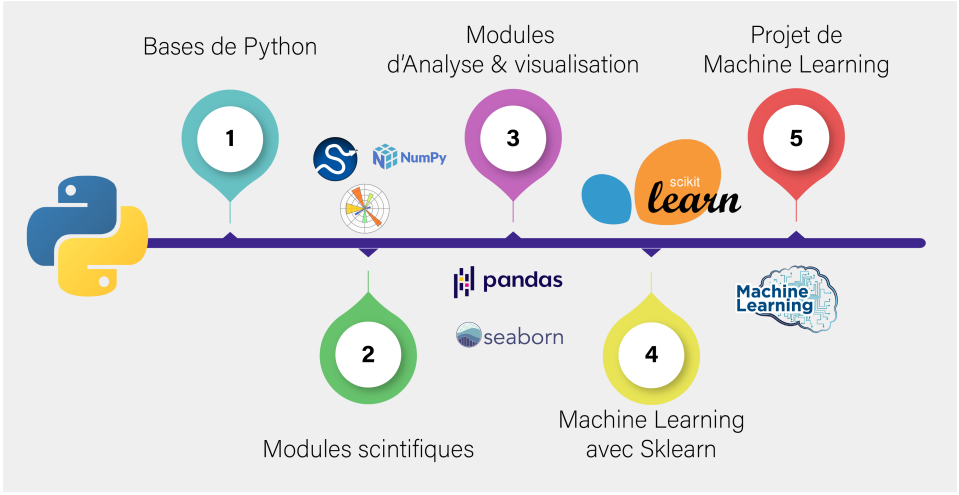


Table of Contents

ML as a Process

3 Machine Learning as a Process

1. Define Objectives.
2. Understand Data (Data Exploratory Analysis)
3. Feature engineering.
4. Data Prepossessing.
5. Model Building.
6. Model Evaluation.
7. Model Tuning
8. Model Deployment

Define Objectives.

- The Define Objectives step is the first stage of the machine learning process, where you define the problem statement and the objectives of the project.
- It involves understanding the business problem and the available data for analysis, in order to set clear and measurable goals.
- The objectives should be aligned with the business needs and provide value to the organization.
- Examples of objectives include predicting customer churn, optimizing product pricing, or improving customer satisfaction scores.
- The Define Objectives step should answer the following questions:
  - What is the problem we are trying to solve?
  - Why is this problem important to solve?
  - What data do we have to work with?
  - What are the success criteria for the project?

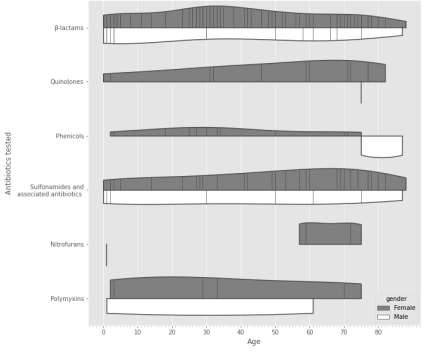
Data Exploratory Analysis.

3 Machine Learning as a Process

In this step, you identify the relevant features (or attributes) that are useful for the machine learning model. This includes selecting, transforming, and creating new features from the available data.

Data Exploratory Analysis

```
1 import pandas as pd
2 import seaborn as sns
3 import matplotlib.pyplot as plt
4
5 df = pd.read_csv("data.csv")
6
7 sns.violinplot(x='age', y='antibiotic',
      hue='gender', split=True,  marker=['o',
       's'],
8                 inner='stick', data=df,cut
      =0,ax=ax)
```



Feature engineering.

3 Machine Learning as a Process

In this step, you identify the relevant features (or attributes) that are useful for the machine learning model. This includes selecting, transforming, and creating new features from the available data.

Code for Feature engineering

```
# Feature selection
selected_features = data[['feature1', 'feature2', 'feature3']]

# Feature transformation
transformed_features = selected_features.apply(lambda x: x**2)

# Feature creation
new_feature = data['feature1'] + data['feature2']
```

Data Prepossessing.

3 Machine Learning as a Process

In this step, you clean, preprocess, and normalize the data to make it ready for analysis. This includes handling missing values, encoding categorical variables, and scaling the features.

Code for Data Prepossessing

```
from sklearn.impute import SimpleImputer
from sklearn.preprocessing import LabelEncoder, MinMaxScaler

# Handle missing values
imputer = SimpleImputer(strategy='mean')
imputed_data = imputer.fit_transform(data)

# Encode categorical variables and Scale the features
label_encoder = LabelEncoder()
scaler = MinMaxScaler()
encoded_data['category'] = label_encoder.fit_transform(data['category'])
scaled_data = scaler.fit_transform(encoded_data)
```

Model Building.

3 Machine Learning as a Process

In this step, you select and train the machine learning model on the prepared data. This includes selecting the appropriate algorithm, splitting the data into training and testing sets, and training the model on the training set.

Code for Model Building

```
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression

# Split data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(scaled_data, target,
    test_size=0.2)

# Train the model
model = LinearRegression()
model.fit(X_train, y_train)
```

Model Evaluation.

3 Machine Learning as a Process

In this step, you optimize the hyperparameters of the model to improve its performance. This includes tuning the regularization parameters, learning rate, and other model-specific parameters.

Code for Model Evaluation

```
from sklearn.model_selection import GridSearchCV

# Define the parameter grid
param_grid = {
    'alpha': [0.01, 0.1, 1, 10],
    'normalize': [True, False]}

# Use grid search to find the best hyperparameters
grid_search = GridSearchCV(estimator=model, param_grid=param_grid, cv=5)
grid_search.fit(X_train, y_train)
best_params = grid_search.best_params_
```

Model Deployment

3 Machine Learning as a Process

In this final step, you deploy the model to a production environment, where it can be used to make predictions on new data.

Code for Model Deployment

```
import joblib

# Save the trained model
joblib.dump(model, 'model.joblib')

# Load the trained model
model = joblib.load('model.joblib')

# Make predictions on new data
new_data = pd.read_csv('new_data.csv')
predictions = model.predict(new_data)
```

Table of Contents

summery

4 Summary

- Python is a popular programming language for data analysis, machine learning, and artificial intelligence.
- Python is easy to learn, has a large user community, and is open-source.
- Python libraries such as NumPy, Pandas, Matplotlib, and Scikit-learn provide powerful tools for data manipulation, analysis, and modeling.
- Machine Learning is a subfield of artificial intelligence that enables systems to learn and improve from experience without being explicitly programmed.
- Machine learning algorithms can be supervised or unsupervised, and can be used for classification, regression, clustering, and other tasks.
- The machine learning process includes several steps, including defining objectives, data exploration, feature engineering, model building, model tuning, model evaluation, and deployment.

Q&A

Thank you for listening!
Your feedback will be highly appreciated!