

Simplified Introduction to Machine Learning with Python

Chapter 1: What is Machine Learning?

Machine learning is a subset of Artificial Intelligence that allows systems to learn and make decisions without being explicitly programmed.

The goal is to build models from sample data to make predictions or decisions based on new data.

Chapter 2: Types of Machine Learning

1. Supervised Learning: Learns from labeled data. Example: Predicting house prices.
2. Unsupervised Learning: Finds patterns in unlabeled data. Example: Customer segmentation.
3. Reinforcement Learning: Learns from actions taken in an environment for rewards or penalties.

Chapter 3: Data Preprocessing

Before training a machine learning model, data needs to be cleaned and transformed. Key preprocessing steps include:

Handling Missing Values:

```
import pandas as pd
df = pd.read_csv('data.csv')
df.fillna(df.mean(), inplace=True) # Fill missing values with column mean
```

Encoding Categorical Variables:

```
from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()
df['category'] = le.fit_transform(df['category'])
```

Feature Scaling:

```
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
```

Simplified Introduction to Machine Learning with Python

```
scaled_data = scaler.fit_transform(df[['feature1', 'feature2']])
```

Chapter 4: Model Training & Evaluation

Scikit-learn is a widely used Python library for machine learning. It provides simple APIs to train and evaluate models.

```
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score

X = df.drop('target', axis=1)
y = df['target']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)

model = LogisticRegression()
model.fit(X_train, y_train)
predictions = model.predict(X_test)
print("Accuracy:", accuracy_score(y_test, predictions))
```