Adaptivity

Introduction to Adaptivity in Machine Learning

Key Concepts

- Adaptivity: ability of a machine learning model to change and evolve based on new data
- Importance of adaptivity: increases the accuracy of machine learning models and enables them to perform well on new and unseen data
- Objectives of the lecture: to explore different types of adaptivity in machine learning

Types of Adaptivity in Machine Learning

Online Learning

- Definition: a form of learning where the model is updated continuously as new data is received
- Advantages: can learn from large, streaming data sets and adapt quickly to changing data
- Examples: stochastic gradient descent, perceptron algorithm
- Applications: fraud detection, recommendation systems, natural language processing

Transfer Learning

- Definition: a technique where knowledge gained from one task is applied to another related task

Types of Adaptivity in Machine Learning (cont.)

- Advantages: enables models to learn from smaller data sets, improves performance on new tasks
- Examples: fine-tuning deep neural networks, domain adaptation
- Applications: image classification, speech recognition, natural language processing

Active Learning

- Definition: a method where the model selects which data to learn from by actively asking for labels from a human expert
- Advantages: reduces the amount of labeled data required for training, enables models to learn from sparse data sets
- Examples: uncertainty sampling, query-by-committee
- Applications: text classification, image classification, speech recognition

Reinforcement Learning

- Definition: a method where the model learns by interacting with an environment and receiving rewards or punishments for its actions

Types of Adaptivity in Machine Learning (cont.)

- Advantages: enables models to learn in a dynamic and changing environment
- Examples: Q-learning, deep reinforcement learning
- Applications: robotics, game playing, autonomous driving

Meta-Learning

- Definition: a method where the model learns how to learn
- Advantages: enables models to adapt to new tasks quickly and efficiently
- Examples: MAML (Model-Agnostic Meta-Learning), meta-reinforcement learning
- Applications: image classification, natural language processing, robotics

Adaptivity in Deep Learning

- Challenges: deep learning models are complex and computationally expensive, making it difficult to adapt to new data quickly
- Opportunities: deep learning models can learn complex and abstract representations, making them well-suited for adaptive learning

- Examples: dynamic neural networks, adversarial training, transfer learning in deep neural networks
- Applications: image recognition, speech recognition, natural language processing

Do you need online learning?

From the River documentation:

You should ask yourself if you need online machine learning. The answer is likely no. Most of the time batch learning does the job just fine. An online approach might fit the bill if:

- You want a model that can learn from new data without having to revisit past data.
- You want a model which is robust to concept drift.
- You want to develop your model in a way that is closer to what occurs in a production context, which is usually event-based.

The Challenge of Dynamic Data in Machine Learning

Data is often considered as static in machine learning courses

- A CSV is provided for analysis
- Students are asked to perform an analysis on it

In practice, data is dynamic

- New data collected continuously
- Models need to be retrained, analyses repeated

Managing Dynamic Data with MLFlow

MLFlow is a framework for managing machine learning workflows

- Keeps track of which model was trained on which data
- Keeps track of software versions, packages used, etc.
- Stores models
- Works with many ML frameworks
- Keeps track of model performance

Benefits of using MLFlow:

- Streamlines the model development process
- Enables efficient troubleshooting of model performance issues

- Facilitates collaboration among data science teams
- Greatly helps with reproducibility

Challenges of Collaboration in Machine Learning

Many industry problems are tackled by teams of data scientists

- Analysis tools like notebooks alone are insufficient
- Keeping track of models created by many collaborators requires specialized tools

Setting up an instance of MLFlow on a server can be technically challenging

- Best for teams to familiarize themselves with MLFlow or similar tools
- MLFlow can be accessed via a web interface for ease of use