FACULTY OF ENGINEERING

Department of Computer and Electrical Engineering

Group Assignment

Course Unit

Selected Topics in Computer Engineering

OPTION: Machine Learning

YEAR 4, SEMESTER I, Date: 5/18/2021

Group One members

Name	Reg.No
OMARA JONATHAN	BU/UG/2017/83
BABISE KEVIN MICHEAL	BU/UG/2017/14
CHELANGAT JOB CHEMOROPUT	BU/UG/2017/

Machine Learning

Machine learning is a form of AI that enables a system to learn from data rather than through explicit programming

Machine learning model is the output generated when you train your machine learning algorithm with data.

You likely interact with machine learning applications without realizing. For example, when you visit an e-commerce site and start viewing products and reading reviews, you're likely presented with other, similar products that you may find interesting

Why has the market of machine learning become so real?

There are six key enablers:

1. Modern processors have become increasingly powerful and increasingly dense. The density to performance ratio has improved dramatically.
2. The cost of storing and managing large amounts of data has been dramatically lowered. In addition, new storage innovations have led to faster performance and the ability to analyze vastly larger data sets.
3. The ability to distribute compute processing across clusters of computers has dramatically improved the ability to analyze complex data in record time.
4. There are more commercial data sets available to support analytics, including weather data, social media data, and medical data sets. Many of these are available as cloud services and well-defined Application Programming Interfaces (APIs).
5. Machine learning algorithms have been made available through open-source communities with large user bases. Therefore, there are more resources, frameworks, and libraries that have made development easier.
6. Visualization has gotten more consumable. You don't need to be a data scientist to interpret results, making use of machine learning broader within many industries.

Approaches to Machine Learning

1. Supervised learning

This is when a machine learning model is trained on labelled data. For example, there could be millions of images of animals and include an explanation of what each animal is and then you can create a machine learning application that distinguishes one animal from another. When the label is continuous, it is a regression; when the data comes from a finite set of values, it known as classification

2. Unsupervised learning

Unsupervised learning is best suited when the problem requires a massive amount of data that is unlabeled. These algorithms are able to classify the data based on the patterns or clusters it finds.

3. Reinforcement learning

Reinforcement learning is a behavioral learning model. The algorithm receives feedback from the analysis of the data so the user is guided to the best outcome. Reinforcement learning differs from other types of supervised learning because the system isn't trained with the sample data set. Rather, the system learns through trial and error.

4. Neural networks and deep learning

Deep learning is a specific method of machine learning that incorporates neural networks in successive layers in order to learn from data in an iterative manner. Deep learning is especially useful when you're trying to learn patterns from unstructured data.

Applying Machine Learning to Business Needs

How companies are beginning to use machine learning techniques to create business differentiation.

1. Understanding why customers are leaving. In order to prevent customer churn, it is critical that you have enough data about the customer's history, his preferences, the services he has purchased in the past, and his complaints
2. Recognizing who has committed a crime through facial recognition
3. Preventing accidents from happening. Machine learning algorithms can be applied to preventive maintenance in a number of ways. For example, a regression algorithm can be used as the foundation for a model that can predict time to failure of a machine

Types of machine learning algorithms

1. Bayesian

Bayesian algorithms allow data scientists to encode prior beliefs about what models should look like, independent of what the data states. These algorithms are especially useful when you don't have massive amounts of data to confidently train a model. A Bayesian algorithm would make sense, for example, if you have prior knowledge to some part of the model and can therefore code that directly. Let's take the case of a medical imaging diagnosis system that looks for lung disorders. If a published journal study estimates the probability of

different lung disorders based on lifestyle, those probabilities can be encoded into the model.

2. Clustering

Clustering is a fairly straightforward technique to understand — objects with similar parameters are grouped together (in a cluster). All objects in a cluster are more similar to each other than objects in other clusters. Clustering is a type of unsupervised learning because the data is not labeled. The algorithm interprets the parameters that make up each item and then groups them accordingly.

3. Decision tree

Decision tree algorithms use a branching structure to illustrate the results of a decision. Decision trees can be used to map the possible outcomes of a decision. Each node of a decision tree represents a possible outcome. Percentages are assigned to nodes based on the likelihood of the outcome occurring.

4. Dimensionality reduction

Dimensionality reduction helps systems remove data that's not useful for analysis. This group of algorithms is used to remove redundant data, outliers, and other non-useful data

5. Instance based

Instance-based algorithms are used when you want to categorize new data points based on similarities to training data. This set of algorithms are sometimes referred to as lazy learners because there is no training phase. Instead, instance-based algorithms simply match new data with training data and categorize the new data points based on similarity to the training data.

6. Neural networks and deep learning

A neural network attempts to mimic the way a human brain approaches problems and uses layers of interconnected units to learn and infer relationships based on observed data

7. Linear regression

Regression algorithms are commonly used for statistical analysis and are key algorithms for use in machine learning. Regression algorithms help analysts model relationships between data points.

8. Regularization to avoid overfitting

Regularization is a technique to modify models to avoid the problem of overfitting. Overfitting occurs when a model is created for a specific data set but will have poor predictive capabilities for a generalized data set.

9. Rule-based machine learning

Rule-based machine learning algorithms use relational rules to describe data. A rule-based system can be contrasted from machine learning systems that create a model that can be

generally applied to all the incoming data. In the abstract, rule based systems are very easy to understand: If X data is inputted, do Y. However, as systems become operationalized, a rule-based approach to machine learning can become very complex

Data used in machine learning

Machine learning algorithms often get the majority of the attention when people discuss machine learning; however, success depends on good data. Your data sources may include both;

I. Traditional systems of record data (such as customer, product, transactional, and financial data) and
II. External data (for example, social media, news stories, weather data, image data, or geospatial data).

In addition, many data structures are critical to analyzing information, including structured and unstructured data.

- **Structured data** refers to data that has a defined length and format e.g. sensor data, financial data, weather data
- **Unstructured data** does not follow any format e.g. Text data, social media data

The Machine Learning Cycle

1. Identify the data: Identifying the relevant data sources is the first step in the cycle. In addition, as you develop your machine learning algorithm, think about expanding the target data to improve the system.
2. Prepare data: Make sure your data is clean, secured, and governed. If you create a machine learning application based on inaccurate data, the application will fail.

3. Select the machine learning algorithm: You may have several machine learning algorithms applicable to your data and business challenge.
4. Train: You need to train the algorithm to create the model. Depending on the type of data and algorithm, the training process may be supervised, unsupervised, or reinforcement learning.
5. Evaluate: Evaluate your models to find the best performing algorithm.
6. Deploy: Machine learning algorithms create models that can be deployed to both cloud and on-premises applications.
7. Predict: After deployment, start making predictions based on new, incoming data.
8. Assess predictions: Assess the validity of your predictions. The information you gather from analyzing the validity of predictions is then fed back into the machine learning cycle to help improve accuracy.

Getting Educated

CognitiveClass.ai https://cognitiveclass.ai

Coursera online learning www.coursera.org/learn/machine-learning

Udacity courses on machine learning www.udacity.com/course/intro-to-machine-learning--ud120

References

1. Machine Learning IBM Limited Edition by Judith Hurwitz and Daniel Kirsch
2. Machine learning: the power and promise of computers that learn by example [online] royalsociety.org/machine-learning