MACHINE LEARNING: A REVIEW

Anshika Sharma[1], Moha Gupta[2]

BCA Student[1], Assistant Professor[2]

KCC Institute of Legal and Higher Education

ABSTRACT: This paper reviews Machine Learning and its aspects in general, why it is important and how it overcomes the real world problems. Although this paper focuses on deductive learning, it at least touches on a great many aspects of Machine Learning in general. In addition, the leading problem of cyber security that could be solved using Machine Learning has been discussed in this review paper. Data from various resources have been used to meet the need of this review paper. Based on the analysis in this review paper, we get to know about the basic advantages and disadvantages of Machine Learning. Also how Machine Learning has proven mandatory to overcome the real world problems and will prove to be helpful regarding the cyber security issues at a greater extent. Because of new computing technologies, machine learning today is not like machine learning of past. By using algorithms to build models that uncover connections, organizations can make better decisions without human intervention. Therefore an individual must learn more about the technologies that are shaping the world we live in.

KEYWORDS: Machine Learning, Deductive Learning, Cyber Security

INTRODUCTION

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it to learn for themselves. The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly.

There are 5 basic steps used to perform a machine learning task:

1. **Collecting data**: Be it the raw data from excel, access, text files etc., this step (gathering past data) forms the foundation of the future learning. The better the variety, density and volume of relevant data, better the learning prospects for the machine becomes.
2. **Preparing the data**: Any analytical process thrives on the quality of the data used. One needs to spend time determining the quality of data and then taking steps for fixing issues such as missing data and treatment of outliers. Exploratory analysis is perhaps one method to study the nuances of the data in details thereby burgeoning the nutritional content of the data.
3. **Training a model**: This step involves choosing the appropriate algorithm and representation of data in the form of the model. The cleaned data is split into two parts – train and test (proportion depending on the prerequisites); the first part (training data) is used for developing the model. The second part (test data), is used as a reference.
4. **Evaluating the model**: To test the accuracy, the second part of the data (holdout / test data) is used. This step determines the precision in the choice of the algorithm based on the outcome. A better test to check accuracy of model is to see its performance on data which was not used at all during model build.
5. **Improving the performance**: This step might involve choosing a different model altogether or introducing more variables to augment the efficiency. That’s why significant amount of time needs to be spent in data collection and preparation.

MACHINE LEARNING IN CYBER SECURITY

We all know, cyber security is the application of technologies, processes and controls to protect systems, networks, programs, devices and data from cyber-attacks. It aims to reduce the risk of cyber-attacks and protect against the unauthorized exploitation of systems, networks and technologies.

Today, it’s impossible to deploy effective cyber security technology without relying heavily on machine learning. At the same time, it’s impossible to effectively deploy machine learning without a comprehensive, rich and complete approach to the underlying data.

Why has machine learning become so critical to cyber security?

Several reasons. With machine learning, cyber security systems can analyze patterns and learn from them to help prevent similar attacks and respond to changing behavior. It can help cyber security teams be more proactive in preventing threats and responding to active attacks in real time. It can reduce the amount of time spent on routine tasks and enable organizations to use their resources more strategically.

In short, machine learning can make cyber security simpler, more proactive, less expensive and far more effective. But it can only do those things if the underlying data that supports the machine learning provides the complete picture of the environment. As they say, garbage in, garbage out.

So far machine learning has proved effective in solving the cyber security problems worldwide. Many tech companies rely on machine learning to overcome the cyber securities threats. For example MICROSOFT.

Microsoft uses its own cyber security platform, Windows Defender Advanced Threat Protection (ATP), for preventative protection, breach detection, automated investigation and response. Windows Defender ATP IS built into Windows 10 devices, automatically updates and employs cloud AI and multiple levels of machine learning algorithms to spot threats.

APPLICATIONS

Machine Learning has many applications in Cyber Security including identifying cyber threats, improving available antivirus software, fighting cyber-crime that also uses AI capabilities, and so on. According to a study conducted by Capgemini Research Institute, AI is necessary for cyber security because hackers are already using it for cyber-attacks.

Top five application of machine learning in cyber security are:

1. Cyber Threat Identification
2. AI-based Antivirus Software
3. User Behavior Modeling
4. Fighting AI Threats
5. Email Monitoring

Sample applications of machine learning:

- **Web search**: ranking page based on what you are most likely to click on.
- **Computational** biology: rational design drugs in the computer based on past experiments.
- **Finance**: decide who to send what credit card offers to. Evaluation of risk on credit offers. How to decide where to invest money.
- **E-commerce**: Predicting customer churn. Whether or not a transaction is fraudulent.
- **Space exploration**: space probes and radio astronomy.
- **Robotics**: how to handle uncertainty in new environments. Autonomous. Self-driving car.
- **Information extraction**: Ask questions over databases across the web.
- **Social networks**: Data on relationships and preferences. Machine learning to extract value from data.
- **Debugging**: Use in computer science problems like debugging. Labor intensive process. Could suggest where the bug could be.

ANALYSIS OF MACHINE LEARNING

Advantages of Machine Learning

Easily identify trends and patterns

- Machine Learning can review large volumes of data and discover specific trends and patterns that would not be apparent to humans.

No human intervention needed (automation)

- Machine Learning does not require physical force i.e., no human intervention is needed.

Continuous Improvement

- ML algorithms gain experience, they keep improving in accuracy and efficiency.

Handling multi-dimensional and multi-variety data

- Machine Learning algorithms are good at handling data that are multi-dimensional and multi-variety, and they can do this in dynamic or uncertain environments.

Disadvantages of Machine Learning

Data Acquisition

- Machine Learning requires massive datasets to train on, and these should be inclusive/unbiased, and of good quality.

Time and Resources

- ML needs enough time to learn and develop enough, to fulfill its purpose with a considerable amount of accuracy and relevancy.
- It also needs massive resources to function.

High error-susceptibility

- Machine Learning is autonomous but highly susceptible to errors.
- It takes time to recognize the source of the issue, and even longer to correct it.

REFRENCES

- Wikipedia
- Quora
- Blogs
- Vlogs
- YouTube videos related to Machine Learning
- PDFs on Machine Learning and Cyber Security
- Article by Syed Muzamil Basha MTech, Dharmendra Singh Rajput PhD, in Deep Learning and Parallel Computing Environment for Bioengineering Systems, 2019
- Article on The Impact of AI on Cyber security by Eddie Segal
- Machine learning in cyber security: A review by Anand Handa, Ashu Sharma, Sandeep K. Shukla
- Machine Learning for Humans, Authors: Vishal Maini, Samer Sabri
- Machine Learning For Dummies®, IBM Limited Edition
- Sites related to Machine Learning and Cyber Security