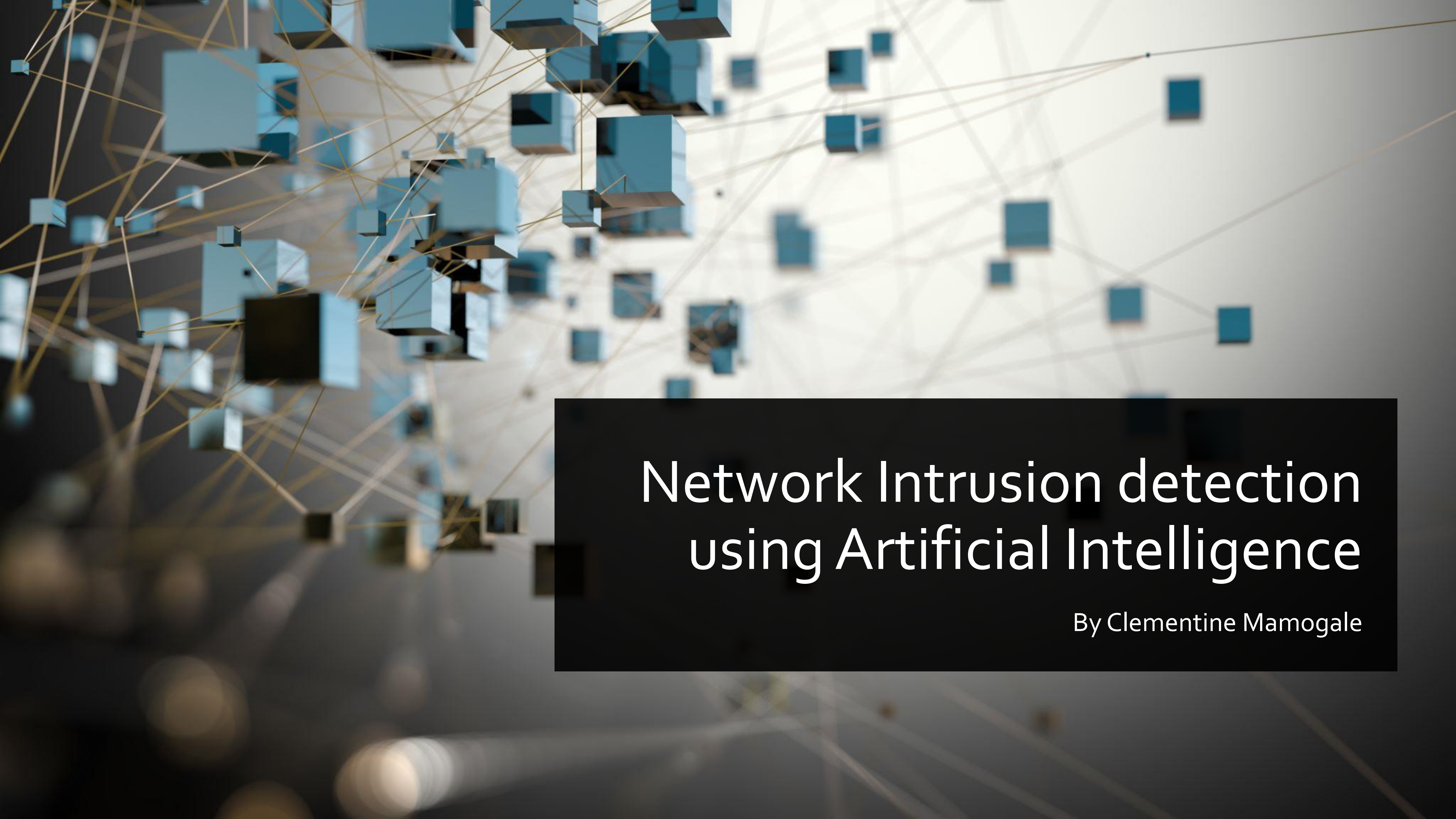


Network Intrusion detection using Artificial Intelligence

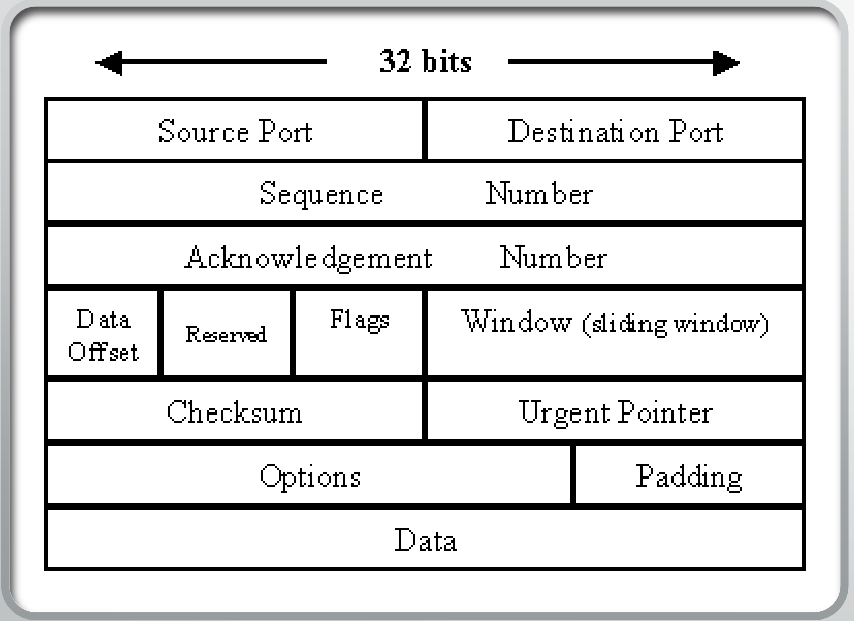
By Clementine Mamogale

Outline

- Introduction
- Research problem
- System overview
- Models
- Results
- Conclusion

Introduction

Network intrusion detection system



32 bits

<table>
<tr><td colspan="3">Source Port</td><td colspan="2">Destination Port</td></tr>
<tr><td colspan="5">Sequence Number</td></tr>
<tr><td colspan="5">Acknowledgement Number</td></tr>
<tr><td>Data Offset</td><td>Reserved</td><td>Flags</td><td colspan="2">Window (sliding window)</td></tr>
<tr><td colspan="3">Checksum</td><td colspan="2">Urgent Pointer</td></tr>
<tr><td colspan="4">Options</td><td>Padding</td></tr>
<tr><td colspan="5">Data</td></tr>
</table>

https://www.techrepublic.com/article/exploring-the-anatomy-of-a-data-packet/

- A security technology that is used to monitor and analyze a network traffic to protect against network-based threats
- Analyze packets by checking the header, content and signature and flag normal or malicious

Research Problem

- Network intrusion detection system (NIDS) is expensive, and only big companies can afford it.
- A NIDS that uses Artificial Intelligence is cheaper, it works better than the traditional NIDS and can be deployed in critical infrastructure

System overview

- Datasets
- Pre-processing
- Model selection
- Training
- Testing
- Classification

Datasets

- **KDD+**
- It is used in many NIDS research papers since it is old, 1999
- Can be used for good baseline of the system
- **CICIDS 2017**
- It is new, it uses modern technologies
- It has datasets that occurred recently

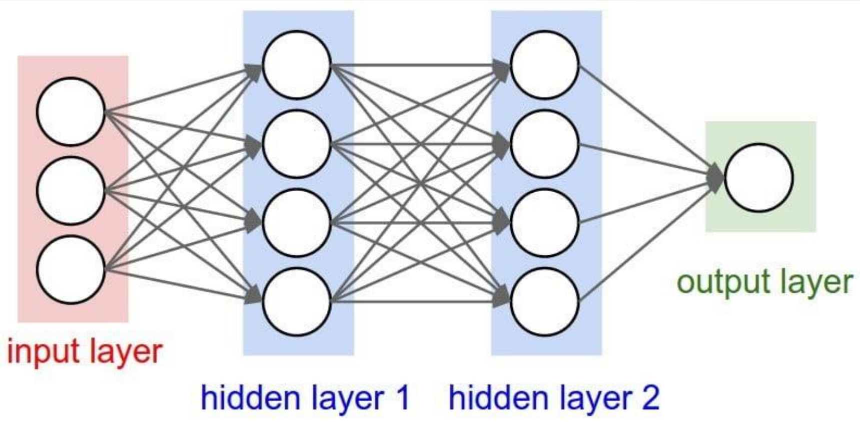
Pre-Processing

- **KDD+ Pipeline**
- Duration
- Protocol type
- Src_bytes
- Dst_bytes
- Labels

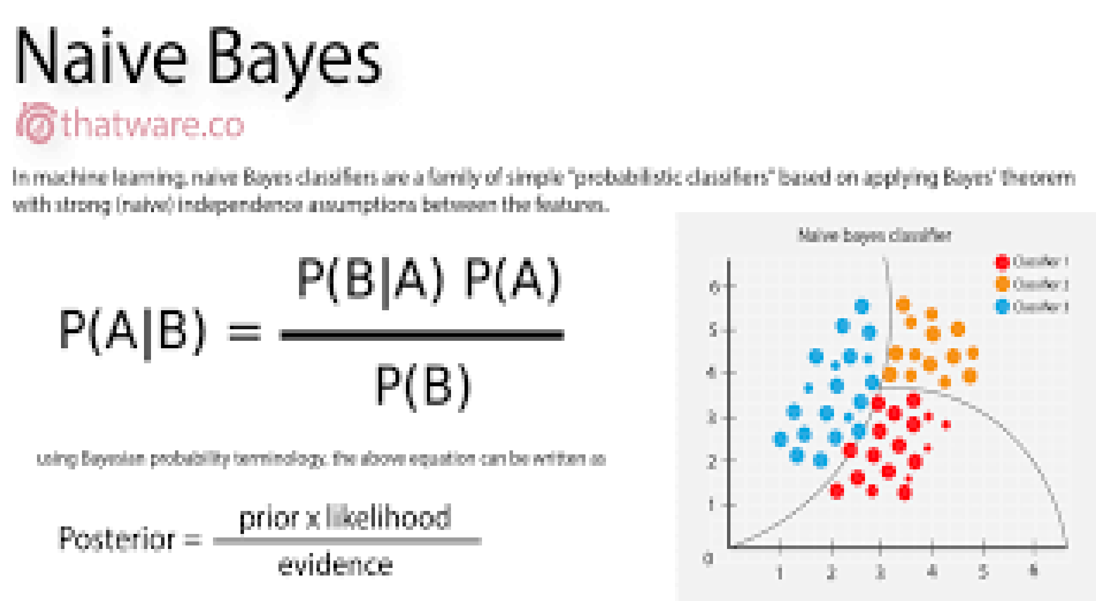
- **CICIDS 2017 Pipeline**
- Flow Duration
- Total Forward
- Total backward
- Forward Packet Length
- Backward Packet Length
- Labels

Model Selection

- Deep Neural Network (DNN)
- Naïve bayes
- Support Vector Machine (SVM)



https://www.bmc.com/blogs/deep-neural-network/



Naive Bayes

thatware.co

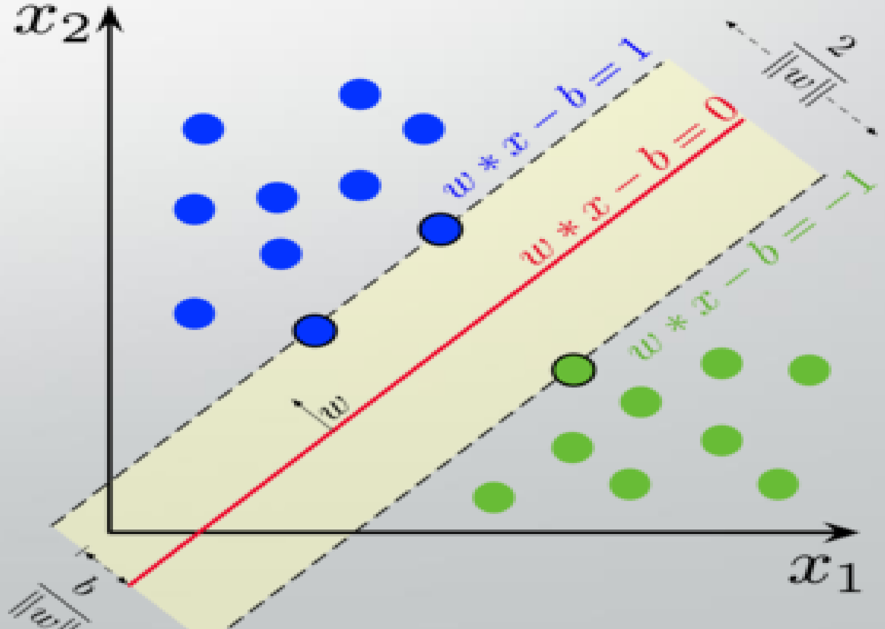
In machine learning, naive Bayes classifiers are a family of simple "probabilistic classifiers" based on applying Bayes' theorem with strong (naive) independence assumptions between the features.

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

using Bayesian probability terminology, the above equation can be written as

$$\text{Posterior} = \frac{\text{prior x likelihood}}{\text{evidence}}$$

https://towardsdatascience.com/introduction-to-na%C3%AFve-bayes-classifier-fa59e3e24aaf



https://en.wikipedia.org/wiki/Support-vector_machine

Results

Model	Dataset	Accuracy
Deep Neural Network (DNN)	KDD+	~92.6
Deep Neural Network (DNN)	CICIDS 2017	N/A
Naïve Bayes	KDD+	~56.2
Naïve Bayes	CICIDS 2017	N/A
Support Vector Machine (SVM)	KDD+	N/A
Support Vector Machine (SVM)	CICIDS 2017	N/A

Conclusion

- I do believe that NIDS can be done using AI, although it might take time to train the model, but once it is trained, it works so well, and it can be so beneficial to many organizations and businesses