

The background of the slide is an abstract visualization of a network. It features numerous blue, three-dimensional cubes of varying sizes, some of which are slightly transparent. These cubes are interconnected by a dense web of thin, golden-yellow lines that crisscross the frame. The overall effect is a sense of a complex, interconnected digital space. The lighting is soft, with a slight gradient from left to right, and some cubes are in sharp focus while others are blurred in the background.

Network Intrusion detection using Artificial Intelligence

By Clementine Mamogale

Outline



Introduction



Research
problem



System
overview



Models

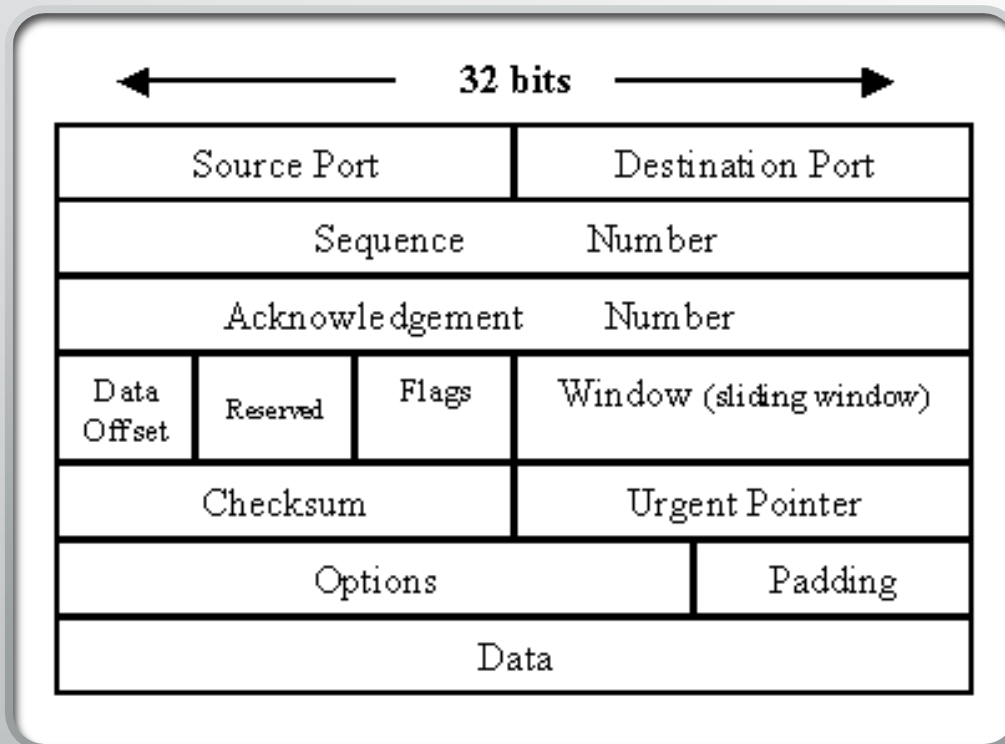


Results



Conclusion

Network intrusion detection system



- A security technology
- monitor and protect network
- Analyze packets - normal or malicious

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

Research Problem

- Network intrusion detection system (NIDS) is expensive
- A NIDS that uses AI is cheaper,
- Works better than the traditional NIDS
- 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

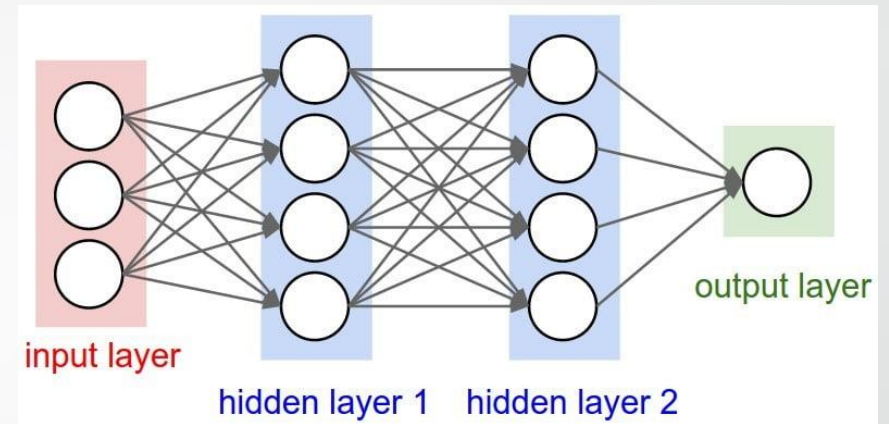
- Duration
- Protocol type
- Src_bytes
- Dst_bytes
- Labels

- 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

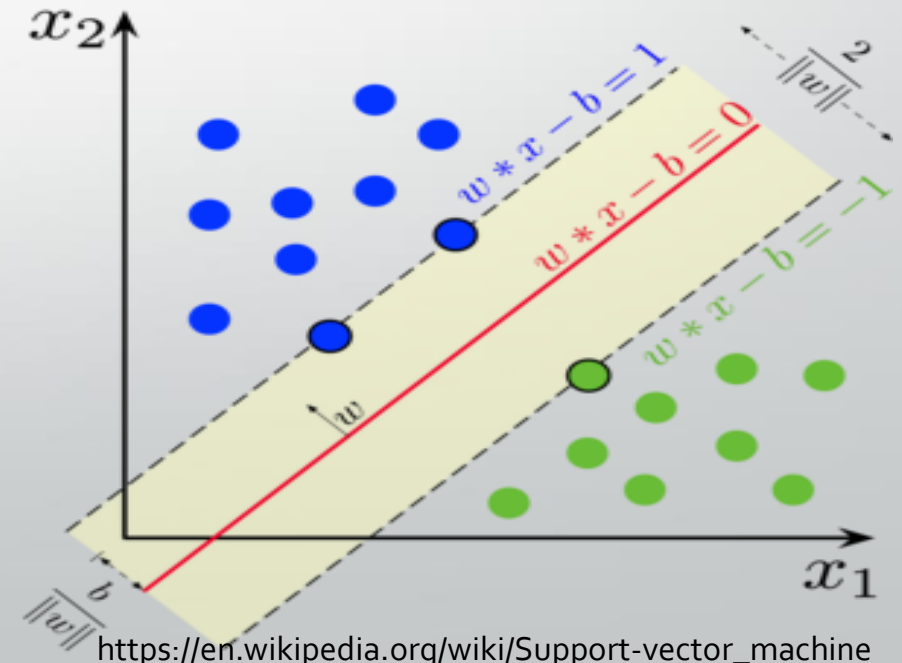
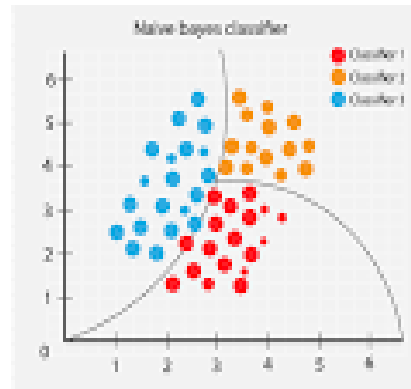
@thatware.co

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

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



https://en.wikipedia.org/wiki/Support_vector_machine

<https://towardsdatascience.com/introduction-to-naive-bayes-classifier-fa59e3e24aaf>

Results

| Model | Dataset | Accuracy |
|------------------------------|-------------|----------|
| Deep Neural Network (DNN) | KDD+ | ~90.92% |
| Deep Neural Network (DNN) | CICIDS 2017 | N/A |
| Naïve Bayes | KDD+ | ~55.6% |
| Naïve Bayes | CICIDS 2017 | N/A |
| Support Vector Machine (SVM) | KDD+ | ~46.38% |
| Support Vector Machine (SVM) | CICIDS 2017 | ~97.48% |

Conclusion

- NIDS can be done using AI
- Takes time to train the model, but...
- Beneficial to many organizations and businesses