Artificial Intelligence (AI) and Machine Learning (ML) for cyber analytics

Ricardo Ramalho

Altice Portugal

Artificial Intelligence (AI)

It’s intelligence—perceiving, synthesizing, and inferring information—demonstrated by machines, as opposed to intelligence displayed by humans or by other animals

Long-term goal: General Intelligence

(the ability to solve an arbitrary problem)

AI fields:

- Reasoning
- Knowledge representation
- Planning
- **Machine Learning (ML)**
- Natural Language Processing (NLP)
- Perception

Why do we need AI?

Every AI field is a specialized “tool”. It’s a not a goal in itself, but is a means

- What does it do?
- What problem does it solve?

Important questions:

- What problems do I want to solve?
- What are the best tools to solve the problem?

- Problem -> Tools to solve it

- Many modern cybersecurity products use AI/ML
  - Intrusion detection
  - Suspicious behaviour

- Machine Learning (ML) important in cybersecurity

Data Science

Data science is the study of data to extract or extrapolate meaningful insights.

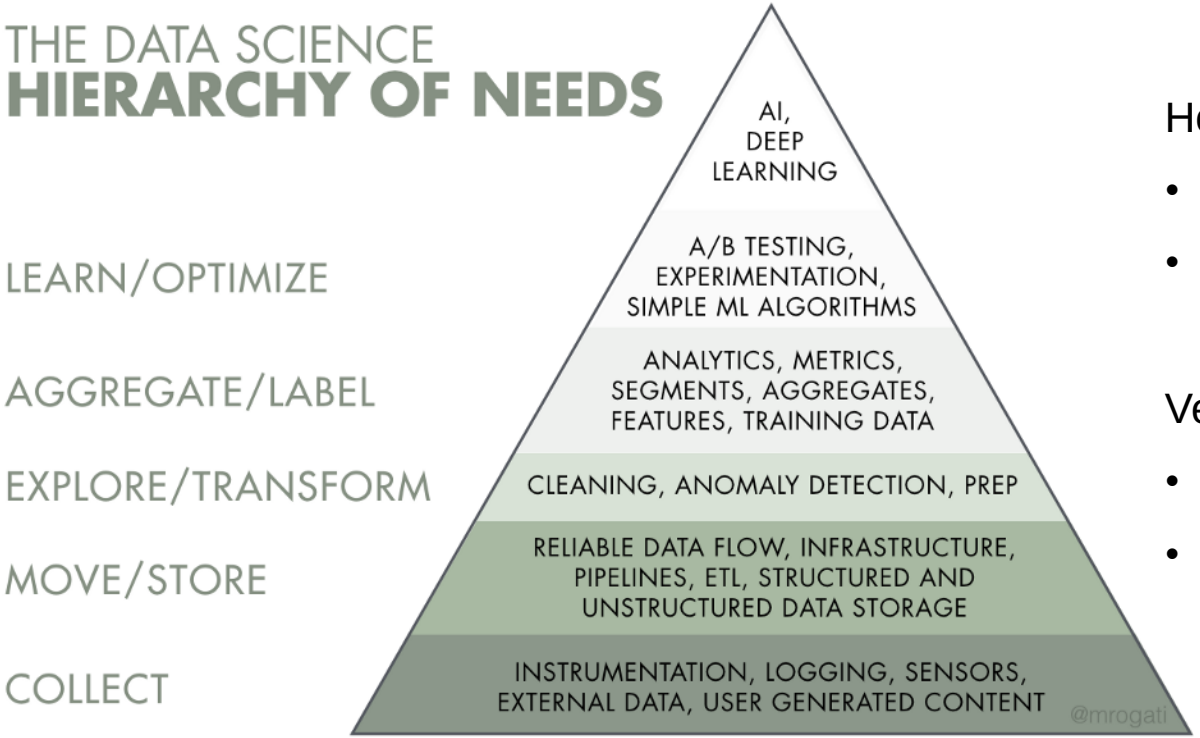
Also:

- multidisciplinary approach
- combines principles and practices from the fields of
  - artificial intelligence (emph. **ML**)
  - mathematics
  - statistics
  - computer engineering
- analyze large amounts of data

Related fields:

- Knowledge Discovery and Data Mining (KDD)
- Analytics

What is needed for Data Science?



Monica Rogati, The AI Hierarchy of needs - Hacker Noon

Horizontal:

- ML libraries
- ETL/Storage tech

Vertical:

- Splunk
- Elastic Search

Data Science in cybersecurity: skillset

Data Scientist

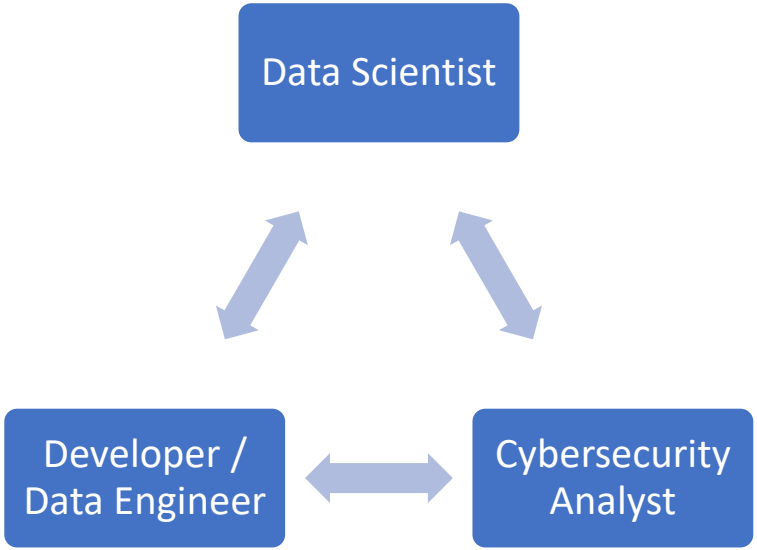
- Data analysis, preparation
- Model creation, training and tuning

Cybersecurity Analyst

- Domain (cybersec) knowledge

Developer / Data Engineer

- Development, Integration
- Data and CI/CD Pipelines



Cybsersecurity uses

Threat hunting

- Data mining / exploration
- Large temporal windows

Create new detectors

- Specific usecase detectors
- Anomaly detection
- Intrusion detection
- Insider abuse

User/Entity Behaviour Analytics (UEBA)

- User and entity (devices) models
- Forensics
- Anomaly detection

Summary

- ‘AI’ is used as a marketing term. It’s a very broad field of research
- Focus on goals/problems first. Then on the tools to solve them (like AI)
- The goal of Data Science is to extract meaningful insights from data
- Implement Data Science Pyramid. Can’t skip steps
- Chose the best / simplest tool for the problem. You will not need AI for everything
- Skillset Trio: Data scientists + Cybersecurity Analysts + Developers **+ Communication**

- Know your users (and devices), know your data, learn from it

Thank you

Ricardo Ramalho
ricardo.g.ramalho@altice.pt