Machine Learning

Exam and AI Hype

Aleksandr Petiushko

ML Research

September 9th, 2023

SOFIA UNIVERSITY

Content

1. Exam information

Content

1. Exam information
2. Exam topics

Content

1. Exam information
2. Exam topics
3. AI/ML/DL “buzzwords” / hype

Exam information

- Exam format: **hand-written on paper**
- Exam duration: **1.5 hrs** (mandatory) + 1 hr (if needed, but no lunch/dinner then)
- Materials allowed to be used: **hand-written notes** + **calculator**
- **No one** is allowed to use either the **phone**, **laptop** or **printed materials**
- Exam parts:
  1. ML Pipeline Design: 60%
  2. Short questions about the course content: 20%
  3. Simple ML-related calculations: 20%

Time for questions

Exam topics: ML pipeline design

Demonstrate the ability to design the ML pipeline for any given problem. It should consist (but not limited to) of the following sub-steps:

- Clear ML task statement
- Data collection strategy
- Data preparation routines
- Model and loss function design
- Success metrics and eval procedure
- Model selection approach

Exam topics: ML concepts

Demonstrate the deep knowledge of the following ML concepts:

- Supervised Learning, types of models (high-level)
- Input feature types and dimensionality
- Empirical vs Structural Risk Minimization
- Overfitting vs Underfitting and methods to avoid them
- Cross-Validation
- Model Selection pipeline and why it is important
- Classification vs Regression
- Classification and Regression loss functions
- Classification quality metrics (including accuracy, precision, and recall)
- Regression quality metrics (including MAE, MSE, and RMSE)
- Binary vs Multi-class Classification
- Micro- vs Macro- Averaging for Multi-class Classification
- L1 (Manhattan) and L2 (Euclidian) norms (distances)
- k-NN Classification and k-NN Regression
- Linear Regression: Ridge, LASSO, and Elastic variants

Exam topics: ML calculations

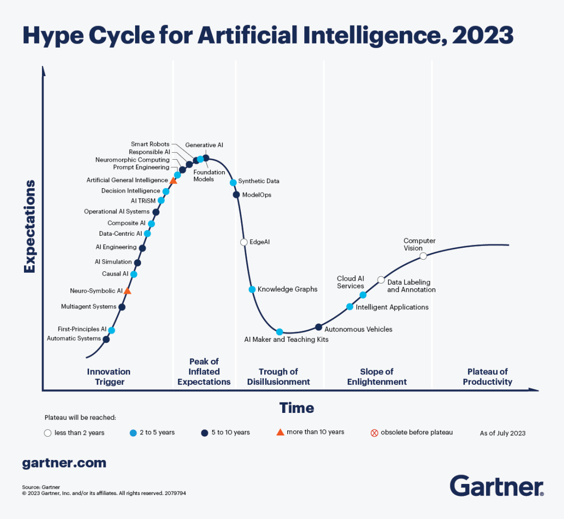
Additionally, to be able to compute auxiliary things like:

- TP, FP, FN, TN
- TPR, FPR, FNR, TNR
- MAE, MSE, RMSE
- Accuracy, Precision, Recall
- Empirical Risk
- L1 (Manhattan) and L2 (Euclidian) norms (distances), and simple equalities/inequalities based on them

AI “buzzwords” requests

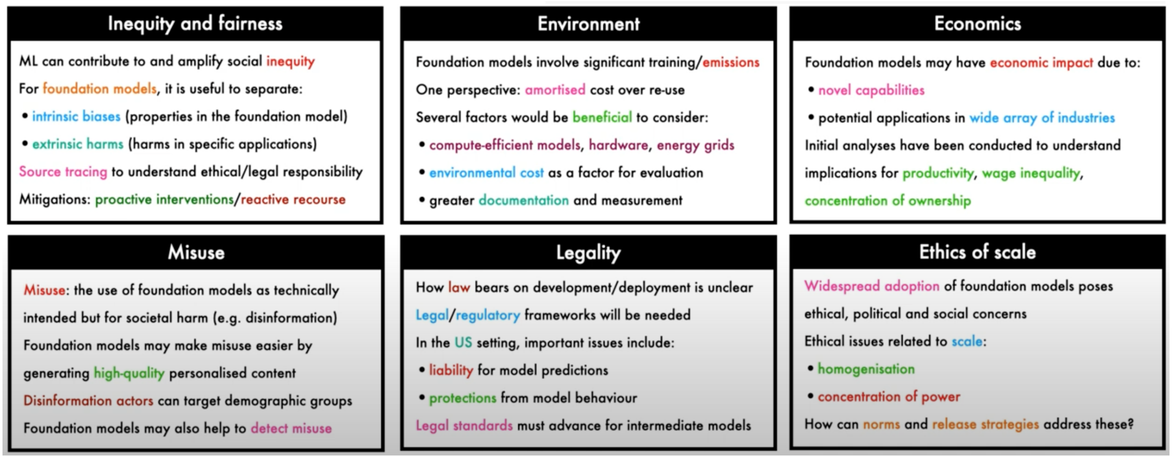
- LLM (GPT / BERT) (application, prompts, mechanisms of training and limited data regime, logical reasoning)
- NLP vs GPT vs Chatbots, Algorithm vs Code, Supervised learning > Semi-supervised learning > Unsupervised
- Generative AI (distributions)
- Embodied AI / Self Driving: what key AI development will be required behind the success of embodied AI (insufficient data of the right form)
- OpenAI, AI Regulations, and Compliance (every app inherits the initial bias, emergent properties)
- AI Art Generator (Stable Diffusion)
- Transfer Learning: Understanding how and when to use transfer learning can save time and computational resources.
- Transformer Architecture
- AI in logistic (NP-hard tasks)
- Different use cases of AI in the real world and how the field is evolving (let’s start with it)

AI Hype Cycle[1]



[1] www.gartner.com

AI ethics and regulations[2]



Inequity and fairness

ML can contribute to and amplify social inequity

For foundation models, it is useful to separate:

- intrinsic biases (properties in the foundation model)
- extrinsic harms (harms in specific applications)

Source tracing to understand ethical/legal responsibility

Mitigations: proactive interventions/reactive recourse

Environment

Foundation models involve significant training/emissions

One perspective: amortised cost over re-use

Several factors would be beneficial to consider:

- compute-efficient models, hardware, energy grids
- environmental cost as a factor for evaluation
- greater documentation and measurement

Economics

Foundation models may have economic impact due to:

- novel capabilities
- potential applications in wide array of industries

Initial analyses have been conducted to understand implications for productivity, wage inequality, concentration of ownership

Misuse

Misuse: the use of foundation models as technically intended but for societal harm (e.g. disinformation)

Foundation models may make misuse easier by generating high-quality personalised content

Disinformation actors can target demographic groups

Foundation models may also help to detect misuse

Legality

How law bears on development/deployment is unclear

Legal/regulatory frameworks will be needed

In the US setting, important issues include:

- liability for model predictions
- protections from model behaviour

Legal standards must advance for intermediate models

Ethics of scale

Widespread adoption of foundation models poses ethical, political and social concerns

Ethical issues related to scale:

- homogenisation
- concentration of power

How can norms and release strategies address these?

[2] www.youtube.com

Transformers

- Overall architecture: `http://jalammar.github.io/illustrated-transformer/`
- Decoder-only variant (GPT): `http://jalammar.github.io/illustrated-gpt2/`

Time for questions

Thank you *all*!