

Innovation



James F. Massa



Fun Career Booster:
Responsible AI that you can Explain, Operate and Govern



You are in the right
place to level up

Fun.ai

Learning Objectives

1. AI lingo level set
2. Fun career booster
3. Why and how we need responsible AI
4. Why and how we need to explain AI
5. Why and how we need to govern AI
6. Which projects to do with AI
7. What are the easiest ways to train an AI model
8. Channel your inner nerd
9. Reinforcement learning quiz
10. Bonus info and call to action

AI lingo level set

Term	Definition
AI	When a computer mimics something a human brain can do—read, write, decide, recommend etc.
Machine Learning (ML)	A subset of AI where a computer program, a “model” is created by training the model on data
Supervised Learning	Learning from data that was labelled by a human
Unsupervised Learning	“Learning” which data is similar and clustering it together
Large Language Model	An ML model trained on enourmous data for more general purpose
Chat GPT	The chat interface for the LLM from OpenAI
Prompt	The command or question input to an LLM
Gen AI	An ML model that can generate new content
Responsible AI	AI that engenders trust and is controlled to avoid bad outcomes
Hallucination	Incorrect or unrealistic output from an LLM
Co-Pilot	AI that works along side you, augmenting your capabilities

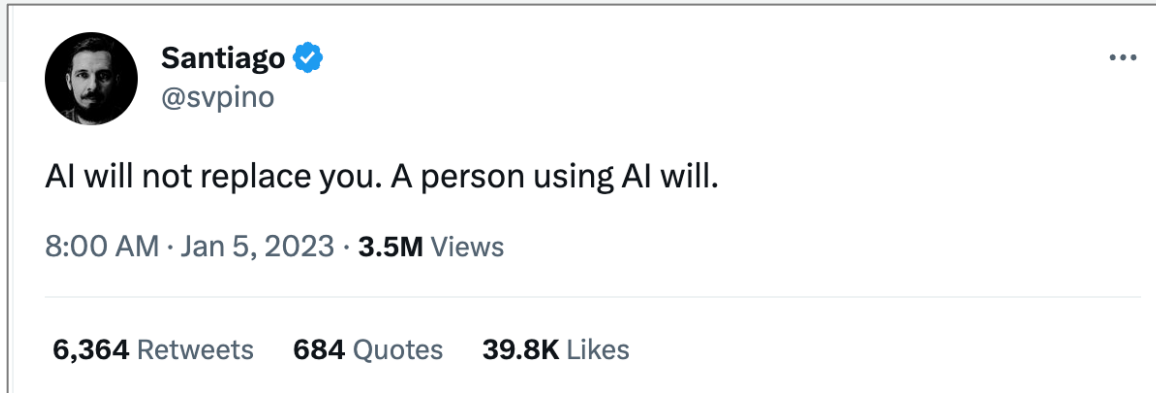




Fun Career Booster

How to embrace AI without being replaced by it

Answers



- AI augmented workers are more productive
- Do more things | Do more complex things
- Inspect more results and approve more decisions
- Consider more factors in decisions

Question

Why and how we need responsible AI

Answers

With large language models comes large responsibility



- **Fairness:** AI can enforce perpetual bias
Aim to reduce bias, not eliminate it, by training on less biased examples
- **Privacy:** LLMs can expose private or copyrighted data
- **Accountability:** AI should be explainable so that all stakeholders can verify performance. AI should be reproducible so we can do retrospectives and make controlled changes.

Question

Why and how we need to explain AI

Answers

Explainability is like a bicycle helmet. You don't need it but crashes hurt less.



- Morality and regulators require transparency
- Show your work, not just the answer
- Visually compare what you found to what you are looking for
- Use explainability and interpretability algorithms



Tough to explain LLMs



Question

Why and how we need to govern AI

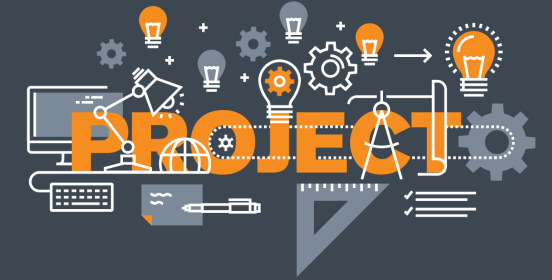
Answers

The eight most comforting words in the English language are: I'm with AI Governance. I'm here to help.

- Who approved this thing?
 - How was it tested and can you explain the results?
 - Are errors and hallucinations tolerable?
- How do you know the data didn't drift?
- Are yesterday's results still appropriate today?

Question

Which projects to do with AI



Answers

- Most people start with anomaly detection
- A chatbot is most people's first LLM project
- Projects with sizeable sustained budgets
- Projects too complex for math rules
- Projects where all the data is available
- Automated approvals and workflows
- Time consuming reading and writing
- Classification, recommendation, and scoring

Question

What are the easiest ways to train the machine



Answers

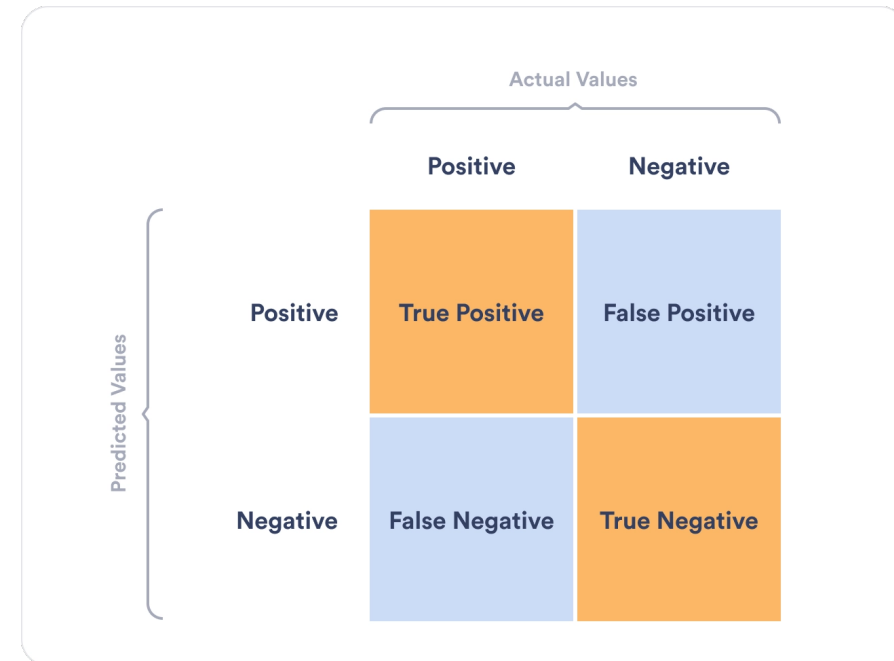
- Instead of training, use pre-trained models, e.g., LLMs
- Anomaly detection with unsupervised learning – no labels
- Leverage already labelled data
- Only train on full and quality data
- Use a lot of data

Channel your inner
nerd question

How to target less false positives and more true positives

Answers

- Understand the confusion matrix
- Prefer Low stakes noise reduction to high stakes problem identification.
- Tune hyper parameters
- Balance Accuracy vs. F1 Score



Reinforcement Learning



QUIZ TIME

A graphic featuring the words "QUIZ TIME" in a stylized, 3D font. Each letter is composed of a grid of small, glowing yellow dots. The letters are colored in a sequence: Q (teal), U (yellow), I (teal), Z (red), T (yellow), I (red), M (teal), and E (yellow). The letters have a slight shadow beneath them, giving them a three-dimensional appearance.

Bonus Question

How to manage a data science project



Answers

- Democratize the data
- Don't cut corners on data or compute
- Integrate data scientists with engineering and business teams
- Hire multiple data scientists
- Reward for going the last mile | production KPIs
- Keep an eye on motivations
- Quadruple year-one estimates

Bonus Question

How to continually improve AI Models





Answers

- KPIs and Confusion Matrix
- Explainability & Hyper Parameters
- Track model performance and drift
- Reinforcement learning
- More balanced training data
- More Features

Call to Action: List three AI points that will you apply



[linkedin.com/in/JamesMassa](https://www.linkedin.com/in/JamesMassa)

- 1.....
- 2.....
- 3.....