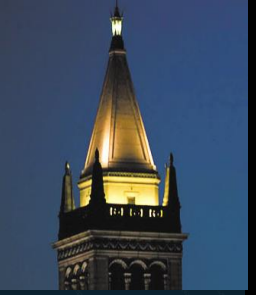


Emerging Trends in AI: Learning, Decisions, and Markets

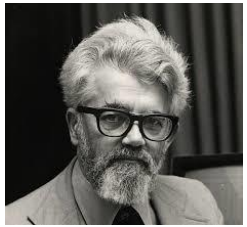
Michael Jordan

University of California, Berkeley

Artificial Intelligence (AI)

- What is it?
- What is it not?
- Where are we currently?
- What can be expected soon?
- What are some implications for business and society?

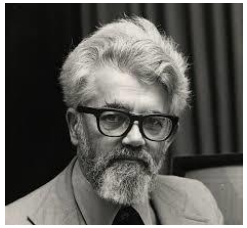
Perspectives on AI*



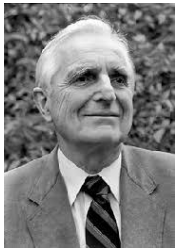
- The classical “human-imitative” perspective
  - cf. AI in the movies, interactive home robotics

*M. I. Jordan, Artificial Intelligence: The Revolution Hasn’t Happened Yet, *Medium*, 2019

Perspectives on AI*

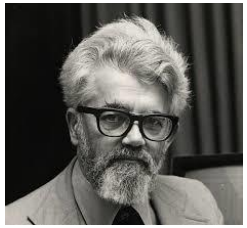


- The classical “human-imitative” perspective
  - cf. AI in the movies, interactive home robotics
- The “intelligence augmentation” (IA) perspective
  - cf. search engines, recommendation systems, natural language translation
  - the system need not be intelligent itself, but it reveals patterns that humans can make use of

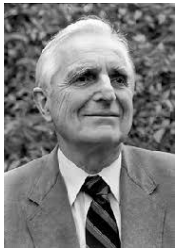


*M. I. Jordan, Artificial Intelligence: The Revolution Hasn’t Happened Yet, *Medium*, 2019

Perspectives on AI*



- The classical “human-imitative” perspective
  - cf. AI in the movies, interactive home robotics
- The “intelligence augmentation” (IA) perspective
  - cf. search engines, recommendation systems, natural language translation
  - the system need not be intelligent itself, but it reveals patterns that humans can make use of



- The “intelligent infrastructure” (II) perspective
  - cf. transportation, intelligent dwellings, urban planning
  - large-scale, distributed collections of data flows and loosely-coupled decisions

*M. I. Jordan, Artificial Intelligence: The Revolution Hasn’t Happened Yet, *Medium*, 2019

Human-Imitative AI: Where Are We?

- Computer vision
  - *Possible*: labeling of objects in visual scenes
  - *Not Yet Possible*: common-sense understanding of visual scenes
- Speech recognition
  - *Possible*: speech-to-text and text-to-speech in a wide range of languages
  - *Not Yet Possible*: common-sense understanding of auditory scenes
- Natural language processing
  - *Possible*: adequate translation and question-answering
  - *Not Yet Possible*: semantic understanding, dialog
- Robotics:
  - *Possible*: industrial programmed robots
  - *Not Yet Possible*: robots that interact with humans and can operate autonomously

What Is Unlikely in our Lifetimes?

- It is unlikely that we will see AI systems that have the intellectual flexibility and creativity of humans
  - AI systems will have limited semantic understanding, and limited ability to cope with complex language (metaphor, irony, etc)
  - AI systems will have limited ability to reason abstractly, finding new abstractions on the fly
  - AI systems will have limited ability to plan in complex environments and adapt their plans on the fly
- They will seem like children who seem to know an amazing number of facts, and can have unusual insights, but nonetheless don’t “understand” very much

What Is Worth Worrying About?

- AI systems that seem smart but actually aren't
  - and make decisions that create massive headaches and disasters
- The loss of jobs and incomes by large numbers of people
  - especially in the service sectors, which have been the main hope for people displaced from classical industrial jobs
  - examples: some call centers, truck drivers, pick-and-place jobs in factories, patent discovery
- Further disparities in wealth and lifestyle if the knowledge of how to build AI systems is not diffused
- The misuse of AI by people with bad intentions

Human-Imitative AI Isn’t the Right Goal

- *To make an overall system behave intelligently, it is neither necessary or sufficient to make each component of the system be intelligent*

- *“Autonomy” shouldn’t be our main goal; rather our goal should be the development of small intelligences that work well with each other and with humans*

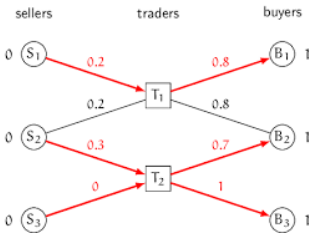
What Intelligent Systems Currently Exist?

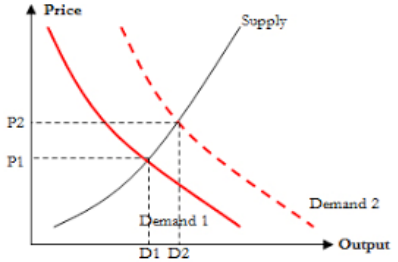
- Brains and Minds

What Intelligent Systems Currently Exist?

- Brains and Minds

- Markets





AI (aka, Machine Learning) Successes

- First Generation (‘90-’00): the backend
  - e.g., fraud detection, search, supply-chain management
- Second Generation (‘00-’10): the human side
  - e.g., recommendation systems, commerce, social media
- Third Generation (‘10-now): pattern recognition
  - e.g., speech recognition, computer vision, translation
- Fourth Generation (emerging): markets
  - not just one agent making a decision or sequence of decisions
  - but a huge interconnected web of data, agents, decisions
  - many new challenges!

Decisions and Markets

- In many problems, a system doesn’t make just a single decision, or a sequence of decisions, but huge numbers of linked decisions in each moment
  - those decisions often interact
- They interact when there is a scarcity of resources
- To manage scarcity of resources at large scale, with huge uncertainty, we need a blend of learning algorithms and markets

Classical Recommendation Systems

- A record is kept of each customer’s purchases
- Customers are “similar” if they buy similar sets of items
- Recommend purchases to new customers by selecting from purchases made by similar customers

Recommendations in the Real World

- Suppose that your company has taken off, and you're making recommendations to large numbers of people

Recommendations in the Real World

- Suppose that your company has taken off, and you're making recommendations to large numbers of people
- Is it OK to recommend the same movie to everyone?
- Is it OK to recommend the same book to everyone?
- Is it OK to recommend the same restaurant to everyone?
- Is it OK to recommend the same street to every driver?
- Is it OK to recommend the same stock purchase to everyone?

The Alternative: Create a Market

- Use data to connect two sides of a market
- Use recommendations on both sides of the market so that the market is efficient
- Intelligence and economic value emerge together
- Issues such as fairness and privacy can also be addressed from this perspective

Example: Music in the Data Age

- More people are making music than ever before, placing it on sites such as SoundCloud
- More people are listening to music than ever before
- But there is no economic value being exchanged between producers and consumers
- And, not surprisingly, most people who make music cannot do it as their full-time job
  - i.e., human happiness is being left on the table

- There do exist companies who make money off of this; they stream data from SoundCloud to listeners, and they make their money … from advertising and subscriptions! ☹

The Alternative: Create a Market

- Use data to provide a dashboard to musicians, letting them learn where their audience is
- The musician can give shows where they have an audience
- And they can make offers to their fans
- I.e., consumers and producers become linked, and value flows: a market is created
  - the company that creates this market profits simply by taking a cut from the transactions

- The company *United Masters* is doing precisely this; see www.unitedmasters.com

The Consequences

- By creating a market based on the data flows, new jobs are created!
- So here's a way that AI can be a job creator, and not (mostly) a job killer
- This can be done in a wide range of other domains, not just music
  - entertainment
  - information services
  - personal services

- The markets-meets-learning approach deals with other problems that a pure learning approach does not
  - e.g., recommendations when there is scarcity

AI = Data + Algorithms + Markets

- Computers are currently gathering huge amounts of data, for and about humans, to be fed into learning algorithms
  - often the goal is to learn to imitate humans
  - a related goal is to provide personalized services to humans
  - but there's a lot of guessing going on about what people want

AI = Data + Algorithms + Markets

- Computers are currently gathering huge amounts of data, for and about humans, to be fed into learning algorithms
  - often the goal is to learn to imitate humans
  - a related goal is to provide personalized services to humans
  - but there's a lot of guessing going on about what people want
- Services are best provided in the context of a market; market design can eliminate much of the guesswork
  - when data flows in a market, the underlying system can learn from that data, so that the market provides better services
  - fairness arises not from providing the same service to everyone, but by allowing individual utilities to be expressed

Markets as Algorithms

- Markets can be viewed as decentralized algorithms
- They accomplish complex tasks like bringing the necessary goods into a city day in and day out
- They are adaptive (accommodating change in physical or social structure), robust (working rain or shine), scalable (working in small villages and big cities), and they can have a very long lifetime
  - indeed, they can work for decades or centuries
  - if we're looking for principles for lifelong adaptation, we should be considering markets as intelligent systems!
- Of course, markets aren't perfect, which simply means that there are research opportunities

Consequences for AI Research

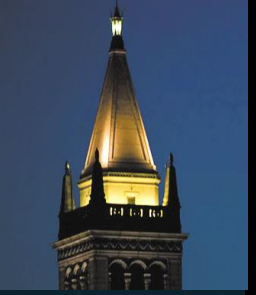
- The focus should be less on autonomy and human-imitative skills
  - there should be less focus on settings in which there is no scarcity and little to gain from interaction
- The focus should be more on federated agents that learn from each other and interact with each other effectively in a world with scarcity and partial information
  - some of these agents are computers and some are humans
  - consider “self-driving cars,” where the federation solves many problems that an autonomous model does not

Consequences for IT Business Models

- Many modern IT companies collect data as part of providing a service on a platform
  - often the value provided by these services is limited
  - so the monetization comes from advertising
  - i.e., many companies are in fact creating markets based on data and learning algorithms, but these markets only link the IT company and the advertisers
- Humans are treated as a product, not as a player in a market
  - the results (ads) are not based on the utility (happiness) of the providers of the data, and does not pay them for their data
- This is broken---humans should be able to participate fully in a market in which their data are being used

Near-Term Technical Challenges

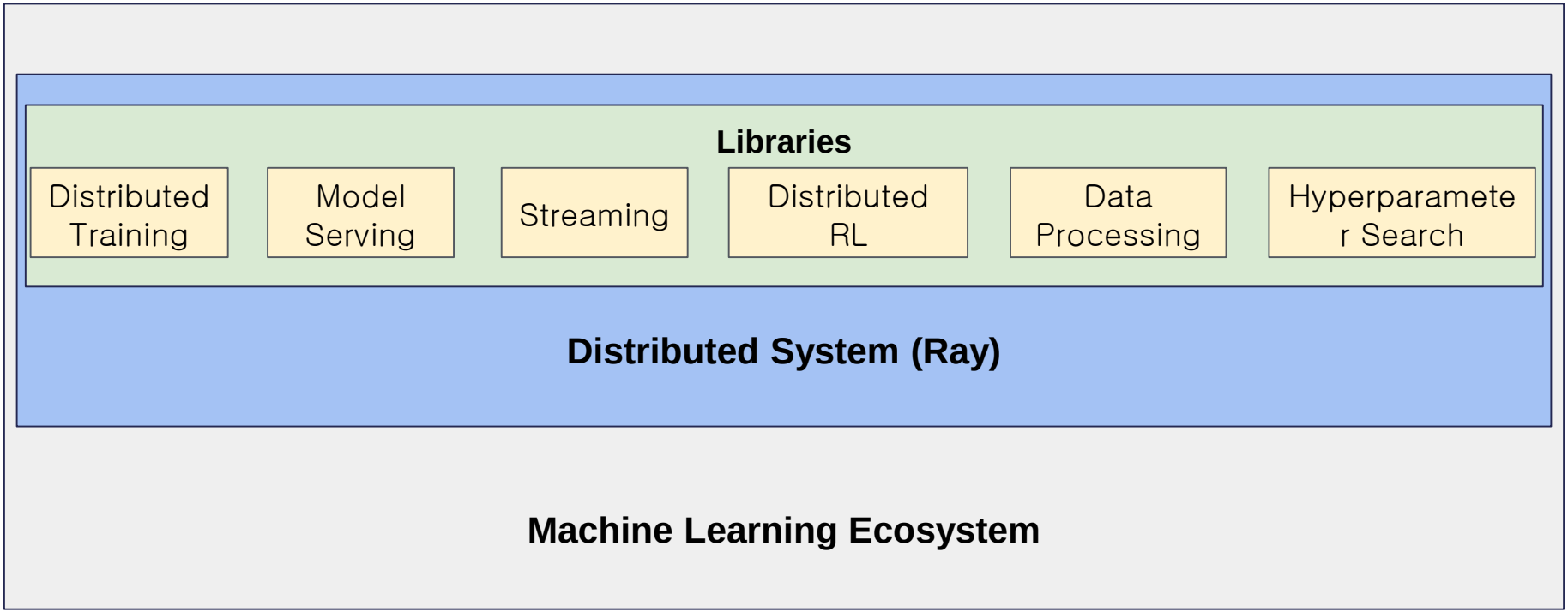
- Error control for multiple decisions
- Systems that create markets
- Designing systems that can provide meaningful, calibrated notions of their uncertainty
- Managing cloud-edge interactions
- Designing systems that can find abstractions quickly
- Provenance in systems that learn and predict
- Designing systems that can explain their decisions
- Finding causes and performing causal reasoning
- Systems that pursue long-term goals, and actively collect data in service of those goals
- Achieving real-time performance goals
- Achieving fairness and diversity
- Robustness in the face of unexpected situations
- Robustness in the face of adversaries
- Sharing data among individuals and organizations
- Protecting privacy and data ownership



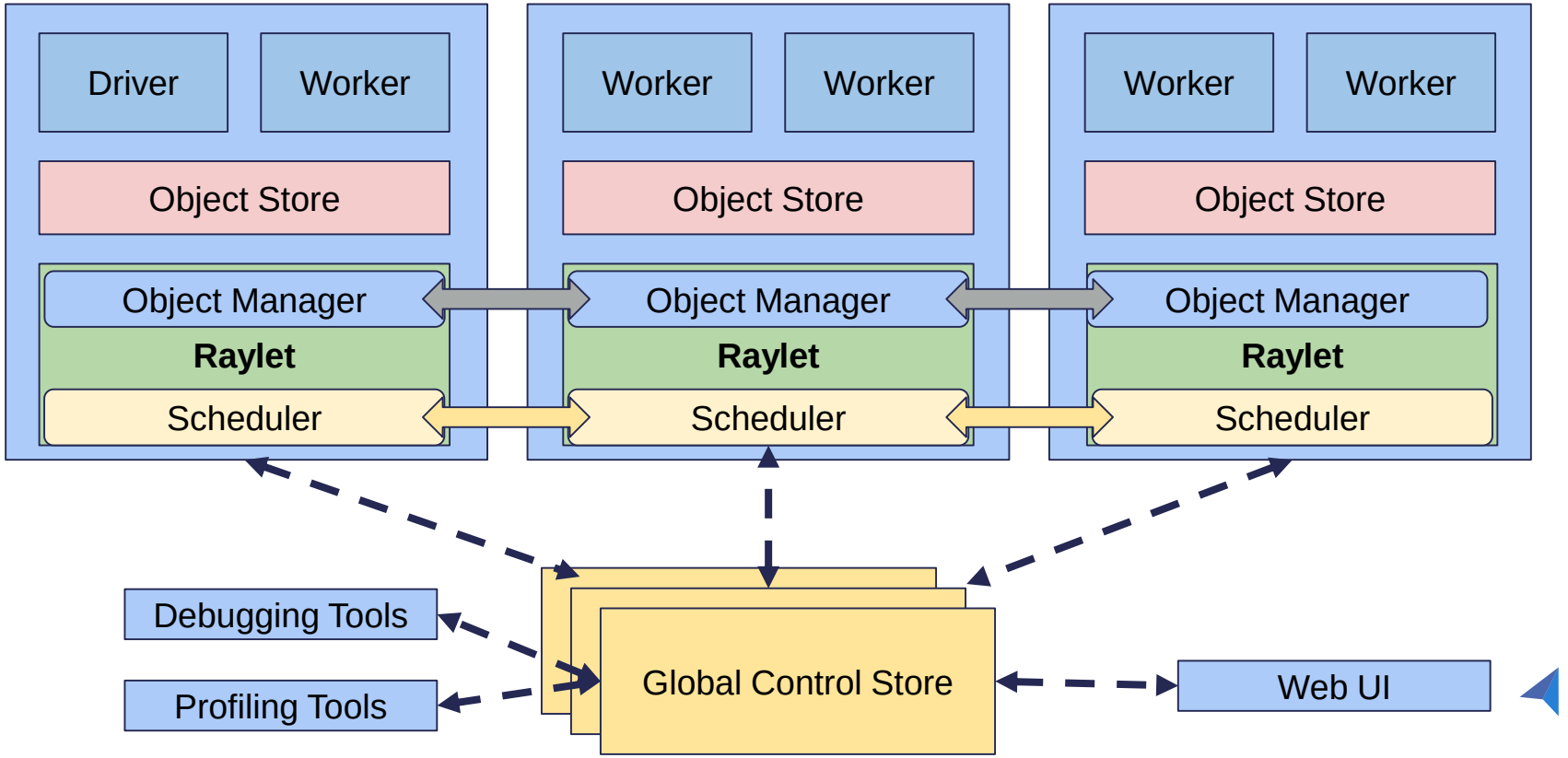
Ray: A Distributed Platform for Emerging Decision-Focused AI Applications

with *P Moritz, R Nishihara, S Wang, A Tumanov, R Liaw, E Liang, and I Stoica*

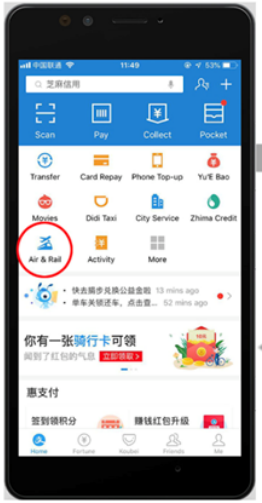
What is Ray?

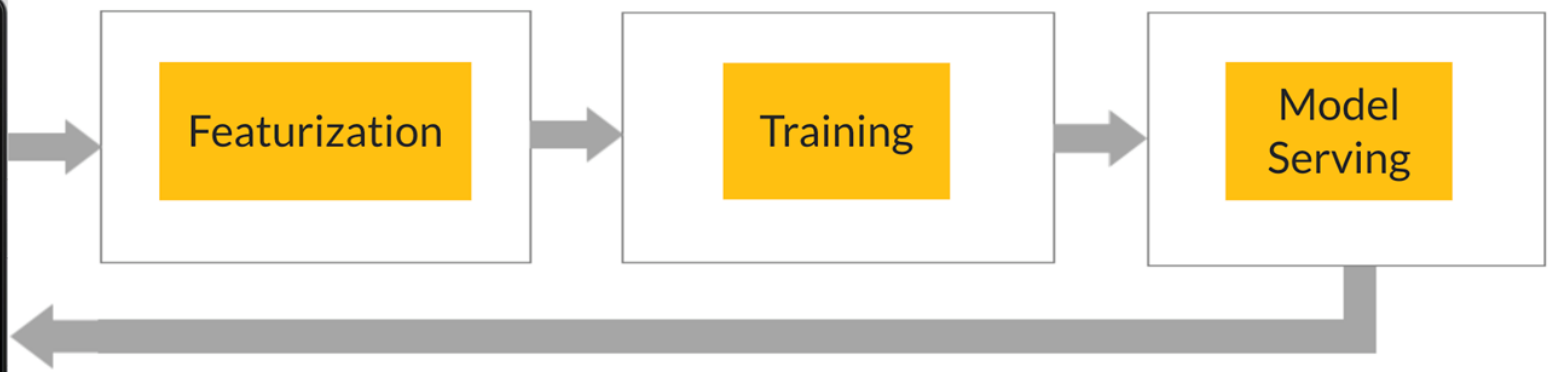

Libraries
Distributed Training
Model Serving
Streaming
Distributed RL
Data Processing
Hyperparamete r Search
Distributed System (Ray)
Machine Learning Ecosystem

Ray Architecture



Example: Online Learning





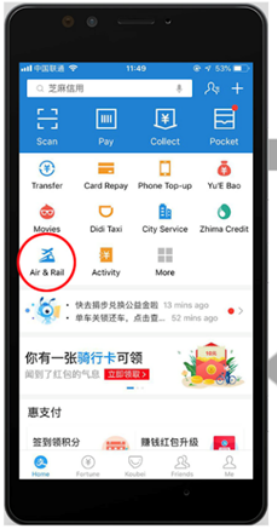
Model updated every **1 day**

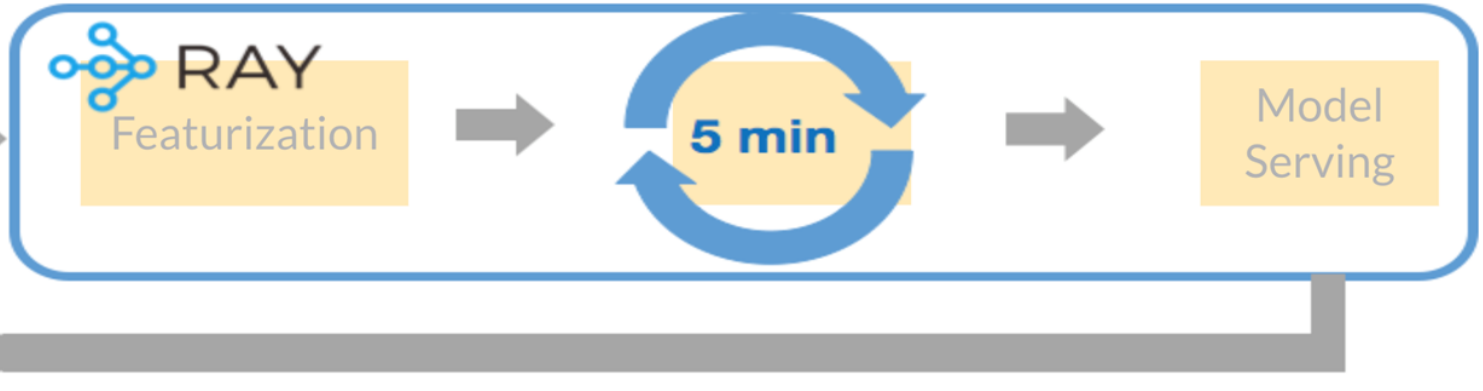
+ 5% CTR

Model updated every **1 hour** (using state-of-the-art solution)

?

Example: Online Learning





Model updated every **1 day**

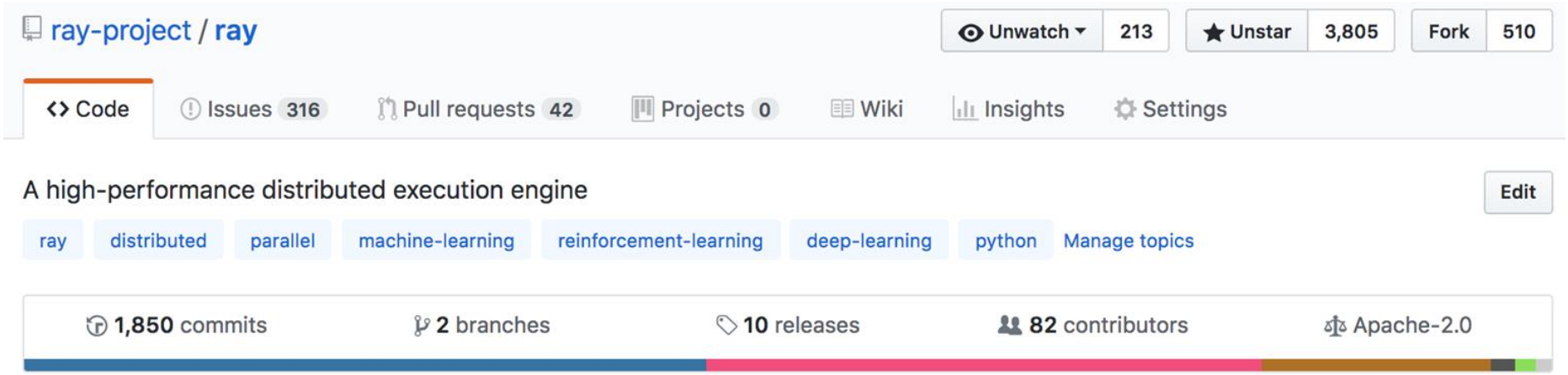
+ 5% CTR

Model updated every **1 hour** (using state-of-the-art solution)

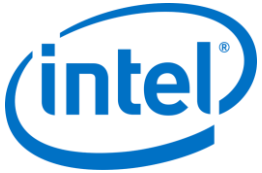
+ 1% CTR

Model updated every **5 min** using **Ray**

Ray is Open Source!



100+ companies using it



“AI” is the Emergence of a New, Human-Centric Engineering Field

- Cf. chemical engineering in the 40s and 50s
  - built on chemistry, fluid mechanics, etc
  - driven by the possibility of building chemical factories
  - new concepts and mathematical principles were needed
- Cf. electrical engineering at the turn of the last century
  - built on electromagnetism, optics, etc
  - new concepts and mathematical principles were needed
- The new field builds on inferential ideas and algorithmic ideas from the past three centuries
  - what’s fundamentally new is the idea of building large-scale systems based on these ideas, using data flows at planetary scale