AI is the new IT: Evidence of the signal

MIT “Intro to machine learning” course enrollment
Artificial intelligence

Machine learning

Neural networks

Deep learning

AI is ...

TECHNICAL DIRECTIONS:

Learn

Represent knowledge

Reason

Interact
1900’s: Santiago Ramón y Cajal
The birth of artificial neural networks

Input layer → Multiple hidden layers process hierarchical features → Output layer

- Identify light/dark pixel value
- Identify edges
- Identify combination of edges
- Identify features
- Identify combination of features

Output "cat"
The deep learning revolution

- Deep Neural Networks
- Large Datasets for Training
- GPU Hardware Accelerators

ImageNet Classification Error

- Sept. 20, 2012

Algorithms
Data
Compute
Specialized ("Narrow") AI works incredibly well

<table>
<thead>
<tr>
<th>Language translation</th>
<th>Speech transcription</th>
<th>Language processing</th>
<th>Visual recognition</th>
</tr>
</thead>
</table>

Think 2019 / DOC ID / September 19, 2019 / © 2019 IBM Corporation
94% of companies believe that AI is the key to competitive advantage

80% of data is either inaccessible, untrusted or unanalyzed

81% do not yet understand the data required for AI

1 in 20 companies have extensively incorporated AI into offerings and processes

60% See compliance as a barrier due to a lack of trust in AI outcomes

65% do not fully trust their own organizations analytics
The evolution of AI

Narrow AI
Emerging

Broad AI
Disruptive and pervasive

We are here

2050 and beyond

General AI
Revolutionary
The evolution of AI

Narrow AI
- Single task, single domain
- Superhuman accuracy and speed for certain tasks
- Deep Learning

Broad AI
- Multi-task, multi-domain
- Learns from less data
- Explainable & Robust Learning & Reasoning

General AI
- Seamless cross-domain learning and reasoning
- Broad autonomy
AI Agenda for the Enterprise

Advancing AI for the enterprise

Trusting AI through fairness, explainability, robustness, transparency

Scaling AI by managing operating and automating its lifecycle

Core AI

Trusted AI

Scalable AI
Advancing the field of AI through grand challenges
<table>
<thead>
<tr>
<th>Project Debater Capabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modeling of human dilemmas and forming principled arguments</td>
</tr>
</tbody>
</table>
Advancing AI

We should subsidize preschools

Stance change

<table>
<thead>
<tr>
<th></th>
<th>Before</th>
<th>After</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>For the motion</td>
<td>79%</td>
<td>62%</td>
<td>-17%</td>
</tr>
<tr>
<td>Against the motion</td>
<td>13%</td>
<td>30%</td>
<td>+17%</td>
</tr>
<tr>
<td>Undecided</td>
<td>8%</td>
<td>8%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Which debater better enriched your knowledge?

- Project Debater: 55%
- Harish: 22%
- Same: 23%
- Undecided: 8%
MIT-IBM Moments in Time: multi-modal action recognition in video
Cognitive Highlights for Sports Events
Unification of learning and reasoning: Neuro-symbolic program induction: Visual question answering (VQA)

1. Scene Parsing (de-rendering)

1. Scene Parsing (de-rendering)

1. Scene Parsing (de-rendering)

2. Question Parsing (program generation)

How many cubes that are behind the cylinder are large?

1. Scene Parsing (de-rendering)

2. Question Parsing (program generation)

How many **cubes** that are behind the **cylinder** are **large**?

3. Program Execution

Neural-Symbolic VQA: Disentangling Reasoning from Vision and Language Understanding

How many blocks are on the right of the three-level tower?

Will the block tower fall if the top block is removed?

What is the shape of the object closest to the large cylinder?

Are there more trees than animals?
By the numbers

$240,000,000 Investment

100 Researchers

10 year Timeline horizon
<table>
<thead>
<tr>
<th>Portfolio to date</th>
<th>186</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposals</td>
<td></td>
</tr>
<tr>
<td>Projects</td>
<td>49</td>
</tr>
<tr>
<td>AI Algorithms</td>
<td></td>
</tr>
<tr>
<td>Physics of AI</td>
<td></td>
</tr>
<tr>
<td>Shared Prosperity with AI</td>
<td></td>
</tr>
<tr>
<td>Security</td>
<td></td>
</tr>
<tr>
<td>Industries</td>
<td>23</td>
</tr>
</tbody>
</table>

Think 2019 / DOC ID / September 19, 2019 / © 2019 IBM Corporation
Advancing AI for the enterprise

**Trusting AI** though fairness, explainability, robustness, transparency

Scaling AI by managing and operating Trusted AI and its lifecycle

Core AI

Trusted AI

Scalable AI
AI is now used in many high-stakes decision making applications
What does it take to trust a decision made by a machine?
<table>
<thead>
<tr>
<th>Is it fair?</th>
<th>Is it easy to understand?</th>
<th>Is it secure?</th>
<th>Is it accountable?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fairness</td>
<td>Explainability</td>
<td>Adversarial Robustness</td>
<td>Transparency</td>
</tr>
</tbody>
</table>

Trusted AI
The pillars of trust

Think 2019 / DOC ID / September 19, 2019 / © 2019 IBM Corporation
Making AI Fair
AI Fairness 360:

- 70+ fairness metrics/checkers
- 10 bias “mitigators”
- industry tutorials

AI Fairness 360 Open Source Toolkit

This extensible open source toolkit can help you examine, report, and mitigate discrimination and bias in machine learning models throughout the AI application lifecycle. Containing over 70 fairness metrics and 10 state-of-the-art bias mitigation algorithms developed by the research community, it is designed to translate algorithmic research from the lab into the actual practice of domains as wide-ranging as finance, human capital management, healthcare, and education. We invite you to use it and improve it.

Not sure what to do first? Start here!

- Read More
  Learn more about fairness and bias mitigation concepts, terminology, and tools before you begin.

- Try a Web Demo
  Step through the process of checking and remediating bias in an interactive web demo that shows a sample of capabilities available in this toolkit.

- Watch a Video
  Watch a video to learn more about AI Fairness 360.

- Read a paper
  Read a paper describing how we designed AI Fairness 360.

- Use Tutorials
  Step through a set of in-depth examples that introduces developers to code that checks and mitigates bias in different industry and application domains.

- Ask a Question
  Join our AIF360 Slack Channel to ask questions, make comments and tell stories about how you use the toolkit.

aif360.mybluemix.net/
Fairness in action: Watson OpenScale

Operationalizes and instruments fairness into enterprise-grade workloads
Making AI Explainable
Algorithm: Contrastive explanations

We deduce that the image is “3” because,
we observe that curves are present,
but a dash is absent.

Trusted AI

Explainability in action: Watson OpenScale

Instruments explanations into enterprise-grade workloads
Making AI Secure
Adversarial Robustness Toolkit

The most comprehensive toolkit for “attacking” and defending AI

https://art-demo.mybluemix.net/
Making AI Transparent
Transparency in action: Factsheets for AI

A transparent reporting mechanism for how an AI system operates and performs
Building more trustworthy AI systems is not only a research question, it’s a business imperative.
AI Agenda for Enterprise

Advancing AI for the enterprise

Trusting AI through fairness, explainability, robustness, transparency

Scaling AI by managing and operating Trusted AI and its lifecycle

Core AI

Trusted AI

Scalable AI
Scaling AI

Accelerating and de-risking AI creation and deployment

Prepare Data  Train  Test & Refine  Deploy  Monitor & Analyze

Operationalizing Trust

Continuous Improvement
Existing approaches leave two important questions unanswered

1. Is my current AI model hurting my KPIs?
2. How do I update my AI model to improve the KPIs?
Improving the model with goal driven active learning

Focus data labelling effort where it is needed the most to improve business KPIs
Advancing AI for the enterprise

Trusting AI through fairness, explainability, robustness, transparency

Scaling AI by managing and operating Trusted AI and its lifecycle

AI Agenda for Enterprise

Core AI

Trusted AI

Scalable AI

Narrow AI
- Single task, single domain
- Superhuman accuracy and speed for certain tasks
- Deep Learning

Broad AI
- Multi-task, multi-domain
- Learns from less data
- Explainable & Robust
- Learning & Reasoning

General AI
- Seamless cross-domain learning and reasoning
- Broad autonomy

IBM Research AI

Think 2019 / DOC ID / September 19, 2019 / © 2019 IBM Corporation