

Guided Tour of Machine Learning in Finance

ML as a foundation of AI - part I

Igor Halperin

NYU Tandon School of Engineering, 2017

Machine Learning as a foundation of AI

- A rational agent should not only perceive its environment, but also **learn** as much as possible from what it perceives

Machine Learning as a foundation of AI

- A rational agent should not only perceive its environment, but also **learn** as much as possible from what it perceives
- Learning = ability to **generalize** using raw data

Machine Learning as a foundation of AI

- A rational agent should not only perceive its environment, but also **learn** as much as possible from what it perceives
- Learning != storing data
- Learning = ability to **generalize** using raw data

Machine Learning as a foundation of AI

- A rational agent should not only perceive its environment, but also **learn** as much as possible from what it perceives
- Learning != storing data
- Learning = ability to **generalize** using raw data

Machine Learning as a foundation of AI

- A rational agent should not only perceive its environment, but also **learn** as much as possible from what it perceives
- Learning != storing data
- Learning = ability to **generalize** using raw data
- Enter **Machine Learning (ML)**: a sub-field of AI that studies perception, learning and action tasks as **algorithms that learn from data**

Machine Learning as a foundation of AI

- A rational agent should not only perceive its environment, but also **learn** as much as possible from what it perceives
- Learning != storing data
- Learning = ability to **generalize** using raw data
- Enter **Machine Learning (ML)**: a sub-field of AI that studies perception, learning and action tasks as **algorithms that learn from data**
- ML modernized such fields as statistics and operational research by putting more emphasis on computer algorithms to statistically estimate complex functions that typically cannot be expressed in closed form

Machine Learning as a foundation of AI

- A rational agent should not only perceive its environment, but also **learn** as much as possible from what it perceives
- Learning != storing data
- Learning = ability to **generalize** using raw data
- Enter **Machine Learning (ML)**: a sub-field of AI that studies perception, learning and action tasks as **algorithms that learn from data**
- ML modernized such fields as statistics and operational research by putting more emphasis on computer algorithms to statistically estimate complex functions that typically cannot be expressed in closed form
- ML is an inter-disciplinary field that benefits from statistical modeling, information theory, physics, computer science, neuroscience, biology

Machine Learning as a foundation of AI

- A rational agent should not only perceive its environment, but also **learn** as much as possible from what it perceives
- Learning != storing data
- Learning = ability to **generalize** using raw data
- Enter **Machine Learning (ML)**: a sub-field of AI that studies perception, learning and action tasks as **algorithms that learn from data**
- ML modernized such fields as statistics and operational research by putting more emphasis on computer algorithms to statistically estimate complex functions that typically cannot be expressed in closed form
- ML is an inter-disciplinary field that benefits from statistical modeling, information theory, physics, computer science, neuroscience, biology
- ML is used in other AI tasks (NLP, knowledge base, AGI)

Machine Learning + AI = Machine Intelligence

Differences between ML and Statistical modeling

Statistical Modeling	Machine Learning
Parametric models that try to “ explain ” the world. The focus is on modeling causality	Non-parametric models that try to “ mimic ” the world rather than “explain” it. Often uses correlations as proxies to causality

Differences between ML and Statistical modeling

Statistical Modeling	Machine Learning
Parametric models that try to “ explain ” the world. The focus is on modeling causality	Non-parametric models that try to “ mimic ” the world rather than “explain” it. Often uses correlations as proxies to causality
Deduce relations for observed quantities by parameter estimation for a pre-specified model of the world	Induce relations between observable quantities, main goal is predictive power

Differences between ML and Statistical modeling

Statistical Modeling	Machine Learning
Parametric models that try to “ explain ” the world. The focus is on modeling causality	Non-parametric models that try to “ mimic ” the world rather than “explain” it. Often uses correlations as proxies to causality
Deduce relations for observed quantities by parameter estimation for a pre-specified model of the world	Induce relations between observable quantities, main goal is predictive power
Small data (1-100 attributes, 100-1000 examples)	Large data (10-100K attributes, 1K-100M examples)

Differences between ML and Statistical modeling

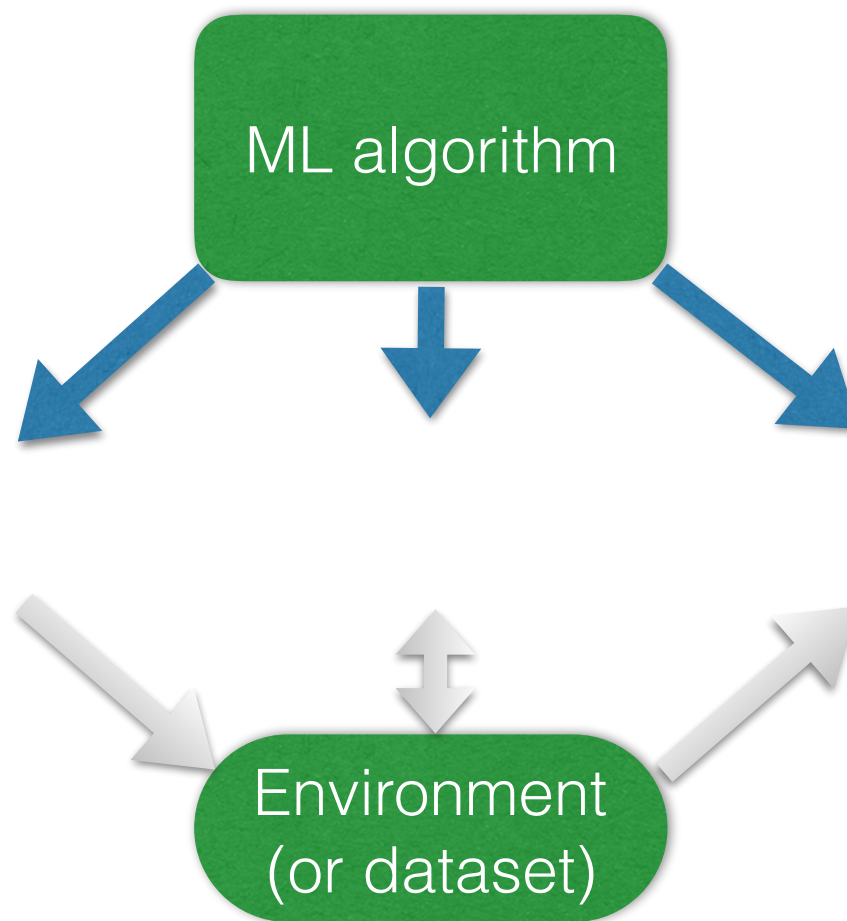
Statistical Modeling	Machine Learning
Parametric models that try to “ explain ” the world. The focus is on modeling causality	Non-parametric models that try to “ mimic ” the world rather than “explain” it. Often uses correlations as proxies to causality
Deduce relations for observed quantities by parameter estimation for a pre-specified model of the world	Induce relations between observable quantities, main goal is predictive power
Small data (1-100 attributes, 100-1000 examples)	Large data (10-100K attributes, 1K-100M examples)
Scalability is typically not the major concern	Scalability is often critical in applications

Differences between ML and Statistical modeling

Statistical Modeling	Machine Learning
Parametric models that try to “ explain ” the world. The focus is on modeling causality	Non-parametric models that try to “ mimic ” the world rather than “explain” it. Often uses correlations as proxies to causality
Deduce relations for observed quantities by parameter estimation for a pre-specified model of the world	Induce relations between observable quantities, main goal is predictive power
Small data (1-100 attributes, 100-1000 examples)	Large data (10-100K attributes, 1K-100M examples)
Scalability is typically not the major concern	Scalability is often critical in applications
Based on a probabilistic approach	Some ML methods are not probabilistic (SVM, neural networks, clustering, etc.)

Machine Learning: core idea

Machine Learning: core idea



Machine Learning: core idea

“A computer program is said to learn from **experience E** with respect to some class of **tasks T** and **performance measure P**, if its performance at tasks in **T**, as measured by **P**, improves with experience **E**.” (Mitchell, 1997)

