

# **Guided Tour of Machine Learning in Finance**

## **ML in Finance vs ML in Tech - part III**

Igor Halperin

NYU Tandon School of Engineering, 2017

# ML in Finance vs ML in Tech

Tasks	ML in Tech	ML in Finance
Big Data?	typically <b>yes</b>	typically <b>no</b>

**Most of data for ML in finance are medium-size, except HFT**

# ML in Finance vs ML in Tech

Tasks	ML in Tech	ML in Finance
Big Data?	typically yes	typically no
Stationary data?	typically <b>yes</b>	most often <b>no</b>

**As most of financial data are non-stationary, collecting more data, even when possible, is not always helpful**

# ML in Finance vs ML in Tech

Tasks	ML in Tech	ML in Finance
Big Data?	typically yes	typically no
Stationary data?	typically yes	most often no
Signal-to-noise ratio	typically <b>low</b>	typically <b>high</b>

**Financial data are typically quite noisy, “true” signals are unobservable!**

# ML in Finance vs ML in Tech

Tasks	ML in Tech	ML in Finance
Big Data?	typically yes	typically no
Stationary data?	typically yes	most often no
Signal-to-noise ratio	typically low	typically high
Action (RL) tasks	<b>Low dimensional state-action space, low uncertainty</b>	<b>High-dimensional state-action space, high uncertainty</b>

- ML in Tech: dimensionality of the state-action space is usually in hundreds. The action space is often discrete (except in robotics). Uncertainty is low to moderate (think self-driving cars!)
- ML in Finance: dimensionality of the state-action space is often in thousands. The action space is usually continuous. Uncertainty is high (think Brexit!)

# ML in Finance vs ML in Tech

Tasks	ML in Tech	ML in Finance
Big Data?	typically yes	typically no
Stationary data?	typically yes	most often no
Signal-to-noise ratio	typically low	typically high
Action (RL) tasks	Low dimensional state-action space, low uncertainty	High-dimensional state-action space, high uncertainty
Interpretability of results	<b>typically, not important, or not the main focus</b>	<b>Typically, either desired or required</b>

**Interpretability of results is:**

- **Desired** for trading
- **Required** for regulation (General Data Protection Regulation, 2018)