Trade and Manage Wealth with Deep Reinforcement Learning

COST - AI in Industry & Finance

ZHAW Winterthur, September 6, 2018

Daniel Egloff, Founder, CEO, Head R&D

FlinkAI

Nobel Laureate Daniel Kahneman on AI

You should replace humans by algorithms whenever possible, and this is really happening. Even when the algorithms don’t do very well, humans do so poorly and are so noisy that just by removing the noise you can do better than people.

FlinkAI

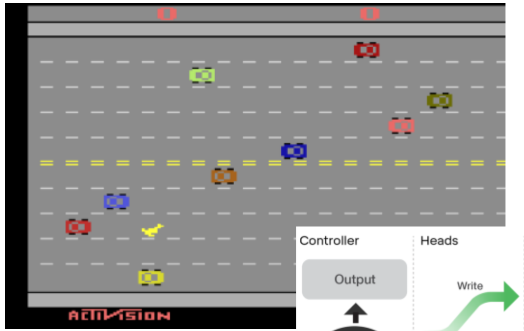
Question

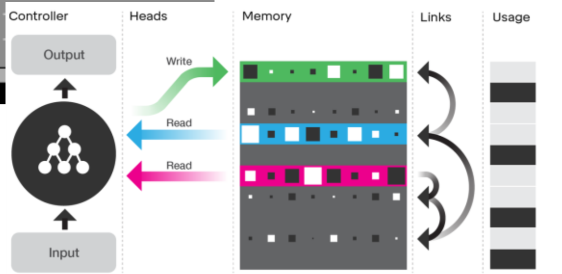
Can modern AI technology replace a PM or trader?

FlinkAI

AI Foundations

- Flink AI builds on several recent innovations in AI and Big Data
  - Deep Reinforcement Learning
  - Differentiable Neural Computer
  - Large scale data streaming infrastructures from eCommerce





What makes DRL Attractive?

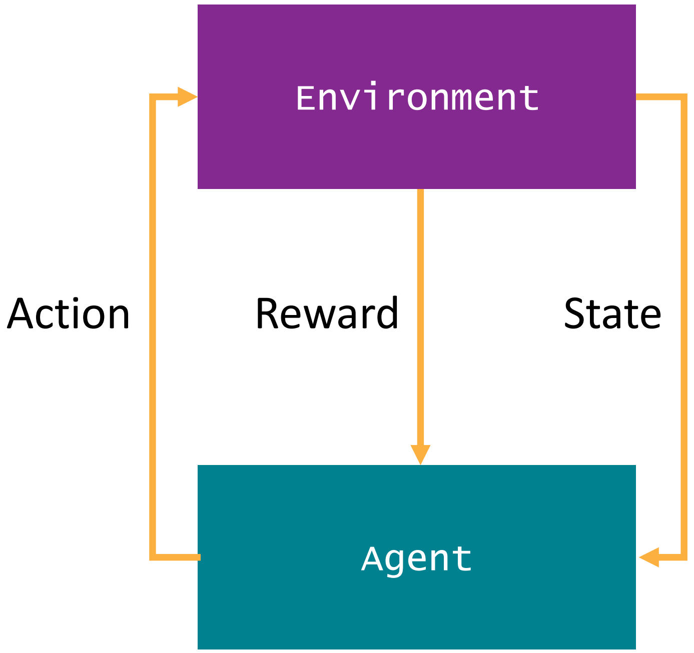
- Construct portfolios and trades from full information
- Feedback loop to improve on good decisions and avoid unsuccessful decisions
- Allows for more realistic modeling of intelligence of a successful PM or trader
- Can work purely data based, model based or combined
- Naturally extends to streaming event data and online learning
- Much easier process

Challenges and Insights

FlinkAI

Reinforcement Learning Setup

Learning a behavioral strategy which maximizes long term sum of rewards by a direct interaction with an unknown and uncertain environment



While not terminal do:

  Agent perceives state s_t

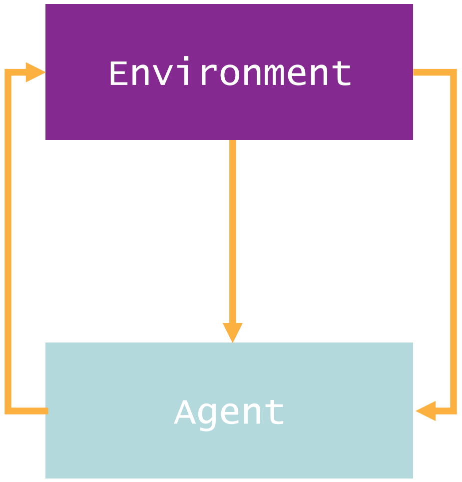
  Agent performs action a_t

  Agent receives reward r_t

  Environment evolves to state s_{t+1}

FlinkAI

RL – State

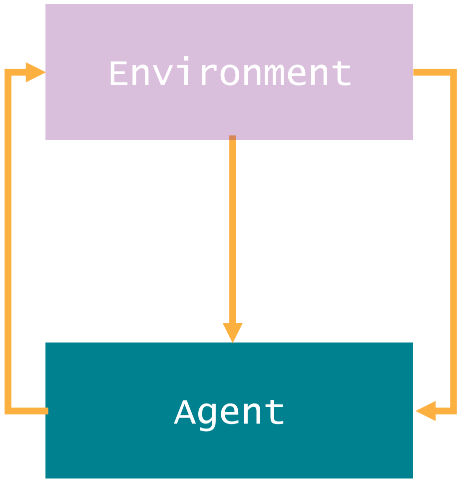


Environment state

- What is the market state?
  - Actual state
  - Market history
  - Market microstructure
  - Large historical state vs recurrent RL
- Which data is required?
  - Price data
  - Limit Order Book (LOB) L1, L2, order flow, trade flow
  - LOB L3 messages
  - Secondary and non standard data
  - Event data versus time clocked data

FlinkAI

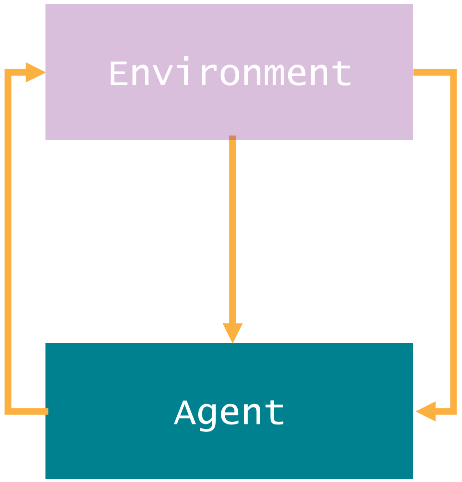
RL – Policy



Agent policy specification

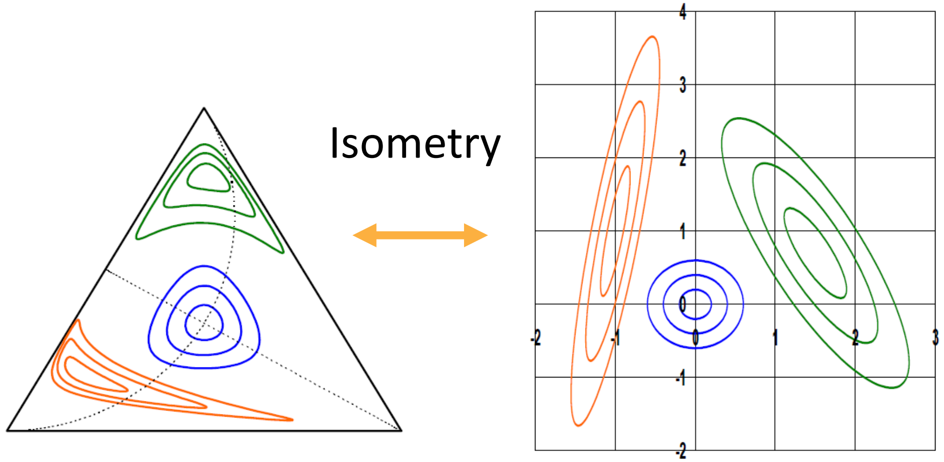
- What is the agent action?
  - Continuous action for percentages of wealth
  - Discrete units of lots to buy/sell
  - Order implementation using market/limit orders
  - Long only vs. long/short
- Long only vs. long/short?
  - Long only agent do not face bankruptcy
  - Short position can lead to bankruptcy

RL – Policy



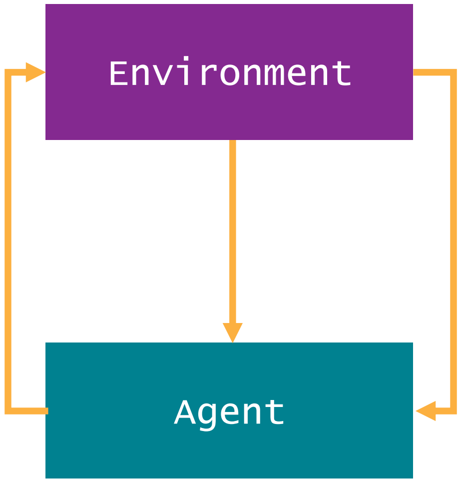
Distributions on the simplex

- Commonly known distributions (Dirichlet, …) not appropriate
- Exploiting less known Hilbert space structure on (open) simplex leading to isometry to Euclidian space (Aitchison)



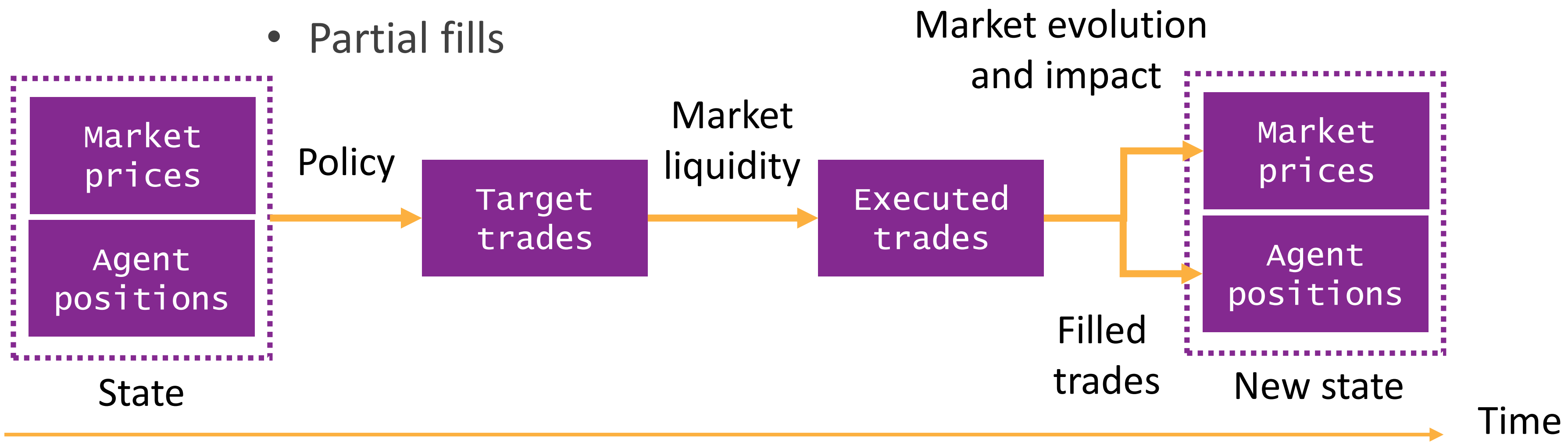
Pull back normal distribution, Student-t, etc.

RL – Interaction



Interaction of agent and environment

- Market evolution, LOB resilience
- Temporal and permanent market impact
- Position change
  - Order cancellations
  - Partial fills



Our 6 Main Challenges

- Sparse trading, learn to use cash and wait for opportunities
- Robustness of RL
- Scaling up RL training
- Handling high resolution event time series data
- Adapting agents to changing markets while not forgetting
- Explaining agent decisions and behavior

Sparse Trading

- Reward modelling, including realistic transaction cost modelling
- Adding risk to give cash a value
- Properly balance risk and reward
- Combining tree search and RL or option framework to learn to postpone trading

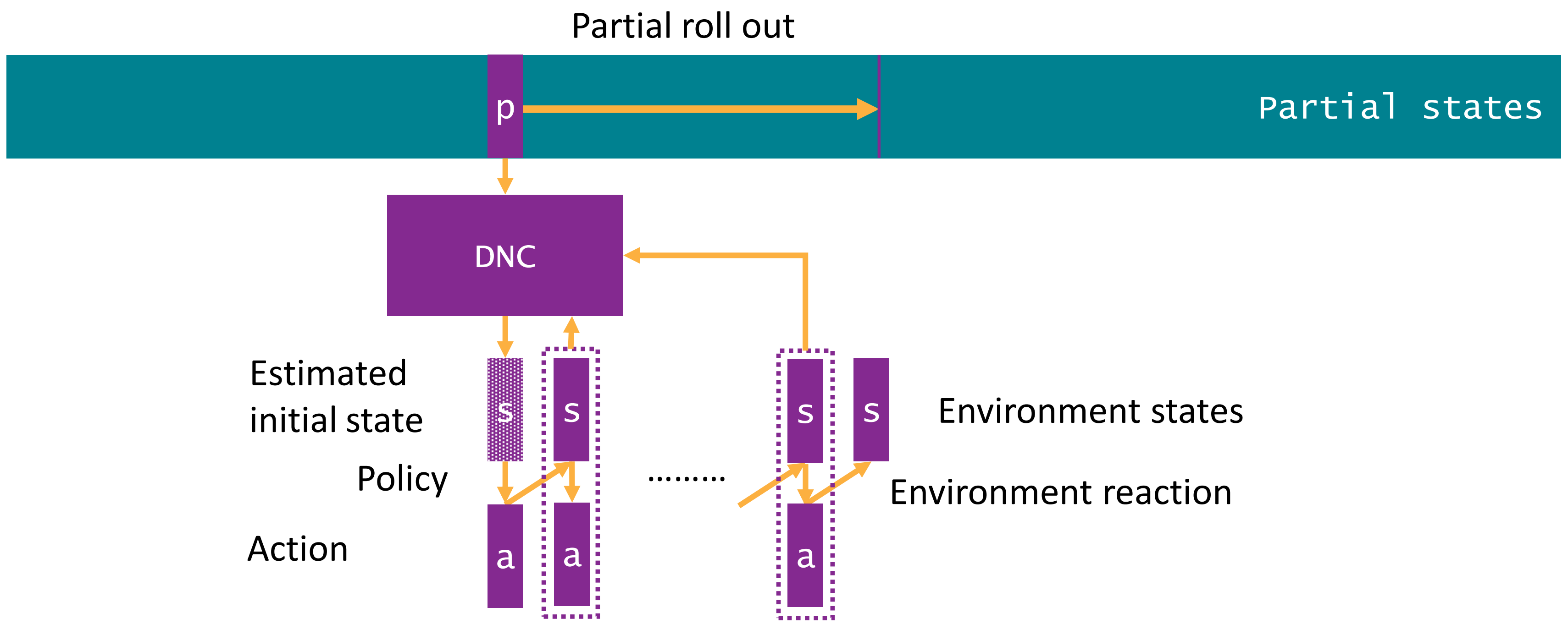
FlinkAI

Robustness

- Reward modelling
- Very long history
- Looking at different scales of time series
- Training on synthesized data, e.g. reconstruct prices from skewed sampling from empirical return distribution

Scaling up RL

- Breaking long episodes into partial episodes with differential memory



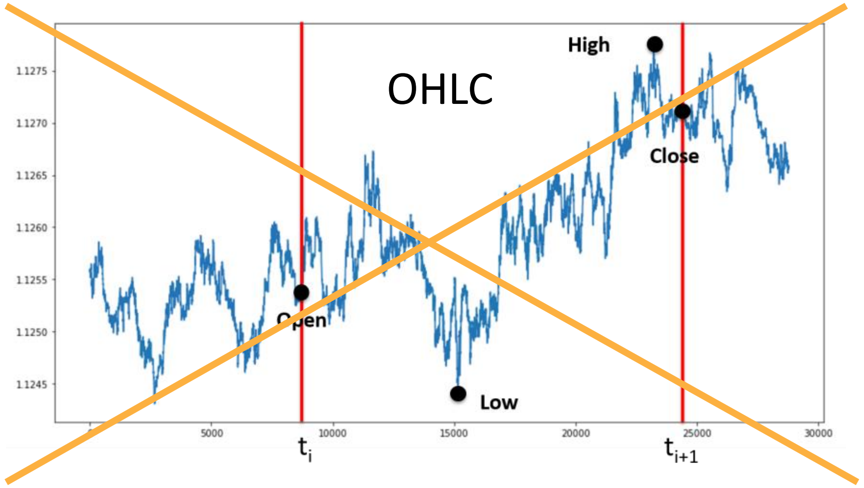
FlinkAI

High Resolution Event TS

Patent pending

- New hybrid RNN – CNN – RNN network topology
- Properly apply convolution over time and cross section
- Cross section should be permutation invariant!
- Convolution at different time frequencies
- Residual NN

OHLC too simplistic

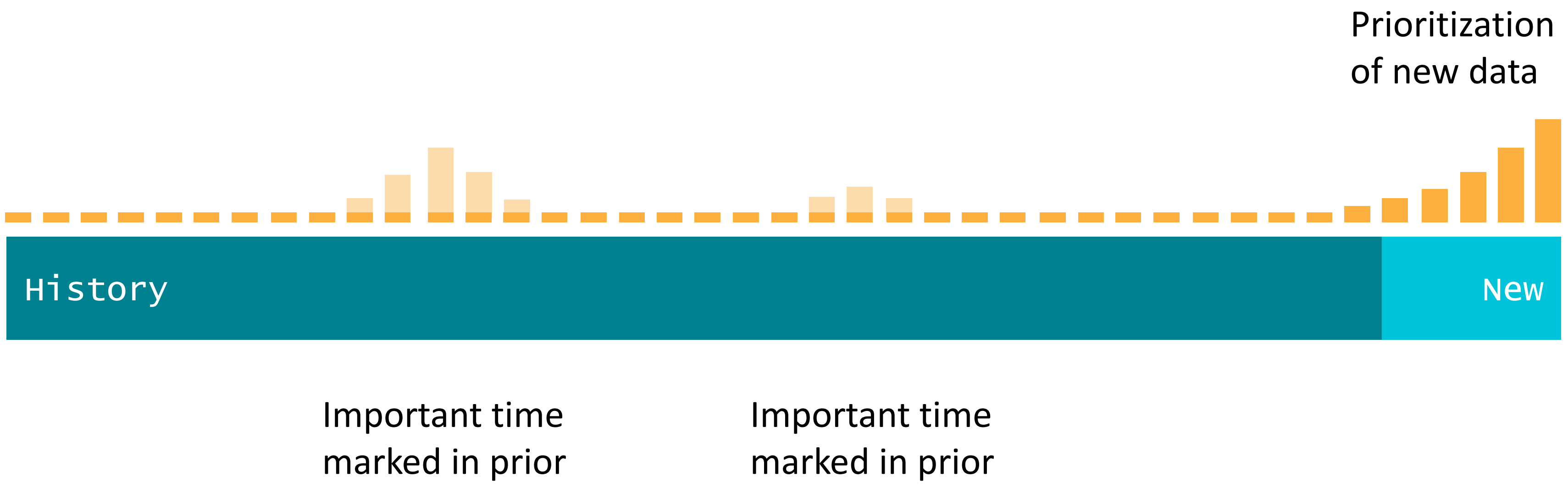


FlinkAI

Adapting while not Forgetting

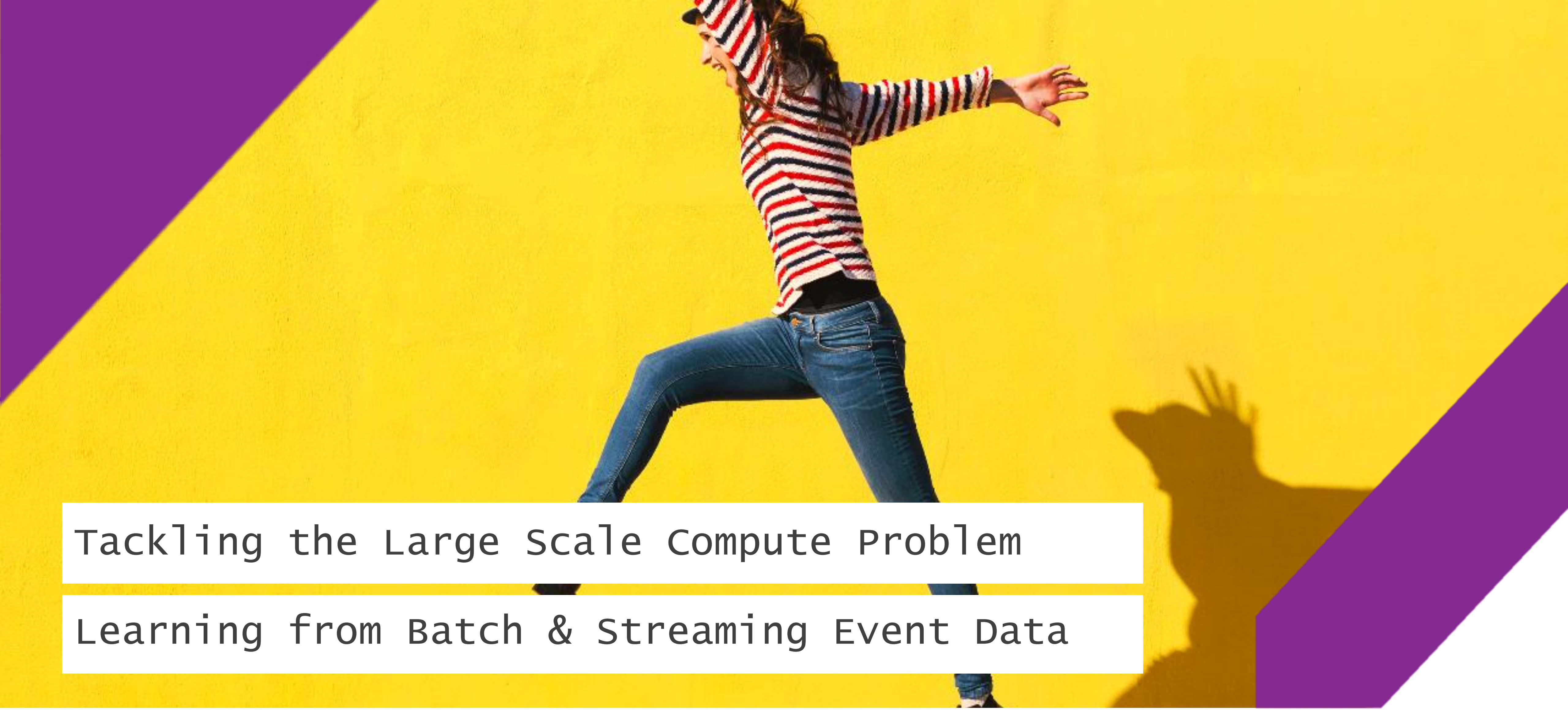
Patent pending

- New attention mechanism relative to prior attention with penalty
- Prior attention reflects “agent style”



Explaining Agent Decisions

- Learning supervised model to explain agent returns
- Compare to different ETF and investment products following a specific investment style
  - Value
  - Growth
  - Momentum
  - Mean reversion



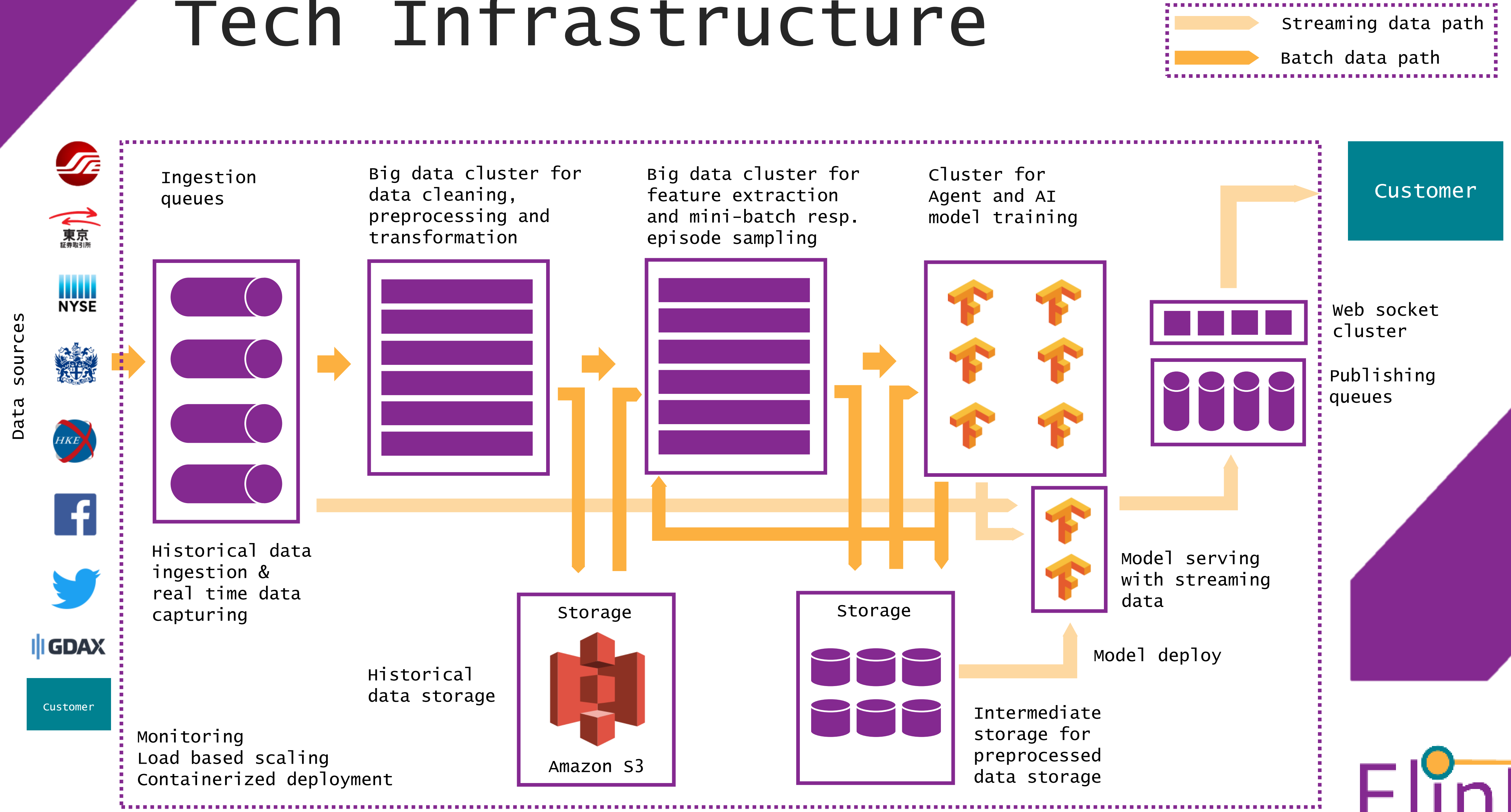
Tackling the Large Scale Compute Problem

Learning from Batch & Streaming Event Data

Large Scale Computing Exercise

- Several compute intense tasks
  - Data preparation (LOB reconstruction, cleaning, …)
  - Feature extraction
  - State labeling
  - RL episode roll out
  - ML model training
- Large amounts of data
  - Storage
  - Network bandwidth and latency
  - Data locality
- Scalability of resources of each task
- Streaming and batch data

Tech Infrastructure



FlinkAI

Tech Infrastructure

- System of multiple clusters
  - Scalability at all levels
  - Fully containerized deployment
  - Redundancy
- Distributed in memory computations whenever possible
- Distributed memory caches to improve IO performance
- Optimizing data locality to reduce network bandwidth
- Unifying batch and streaming data processing at all stages
- Incremental learning from new data and combining with historical data
- Feedback loop from training to serving & training to episode generation
  - Online and incremental learning
  - Reinforcement learning

FlinkAI

Contacts

www.flink.ai | daniel@flink.ai