XAI in Financial Services

Daniele Magazzeni, PhD
AI Research Director
Head of Explainable AI Center of Excellence
J.P. Morgan

AI Research at J.P. Morgan

We have
80 research projects
across our 7 Research Goals

Stakeholders | Financial Domain | Cross-Cutting


J.P.Morgan
Explainable AI
Center of Excellence
AI Research

The Need for Explainable AI in Finance

- Some of the ML models used in finance are complex and non-linear. Hence, the overall functioning of some ML models, and how they arrive at a particular outcome, is not always obvious
- XAI is key for Customers, Regulators, Practitioners, Risk Managers

Key risks

- **Lack of transparency** to internal users and external clients
- Difficulty **in complying with regulatory expectations**, including fair lending and adverse action notification
- Difficulty in **assessing conceptual soundness** of the model during model validation

Types and Scope of Explanations

Global Explanations: give insights on the overall behaviour of the model

Local Explanations: give insights on a specific output of a model

Explanations as **Verification and Robustness Testing**

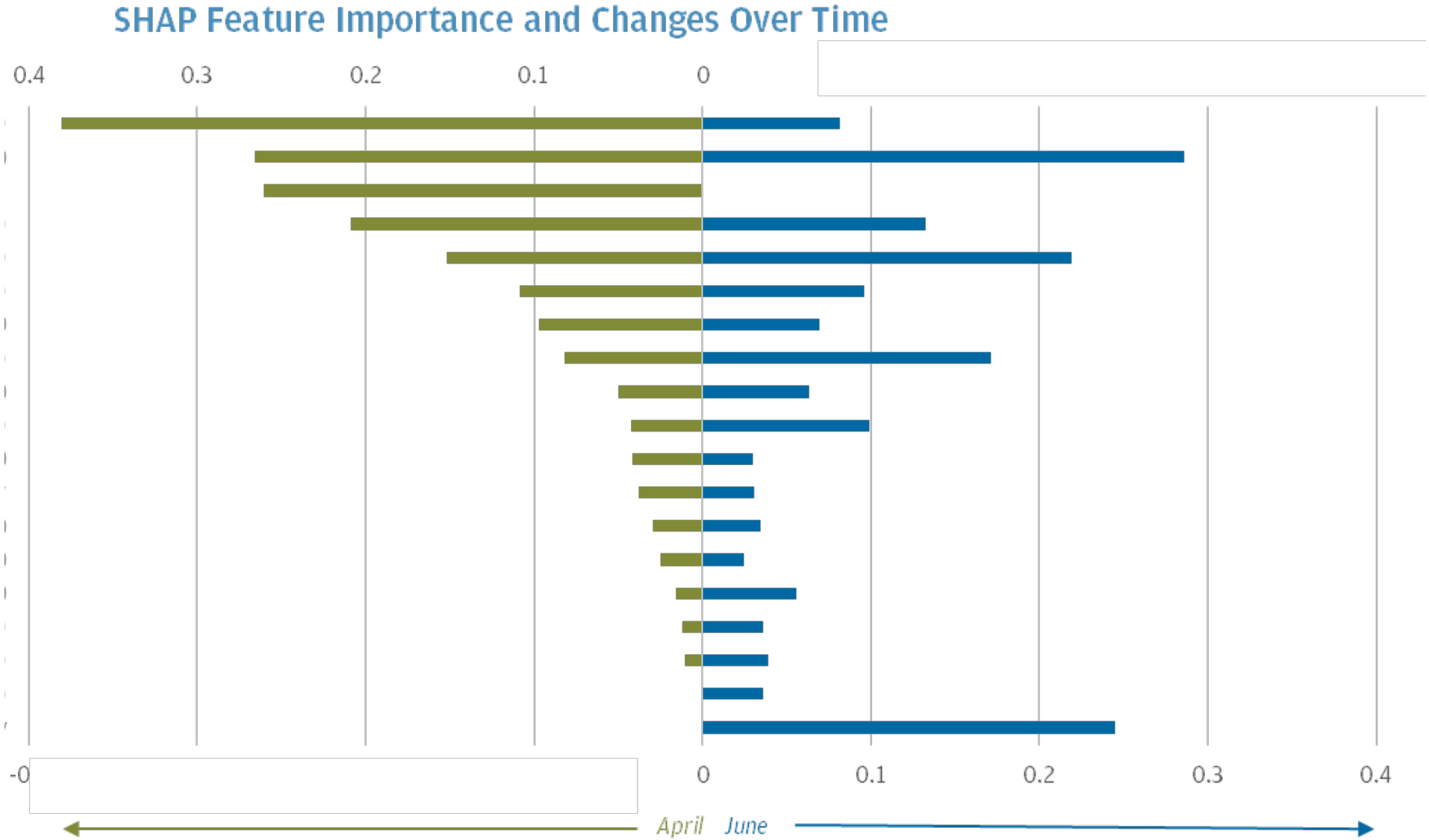
- Target: Risk Management, Regulators, and AI developers
- Questions: *"Is the model compliant with stress test scenarios?"* or *"Is the model consistent with fair lending requirements?"*

Explanations as **Diagnostics**

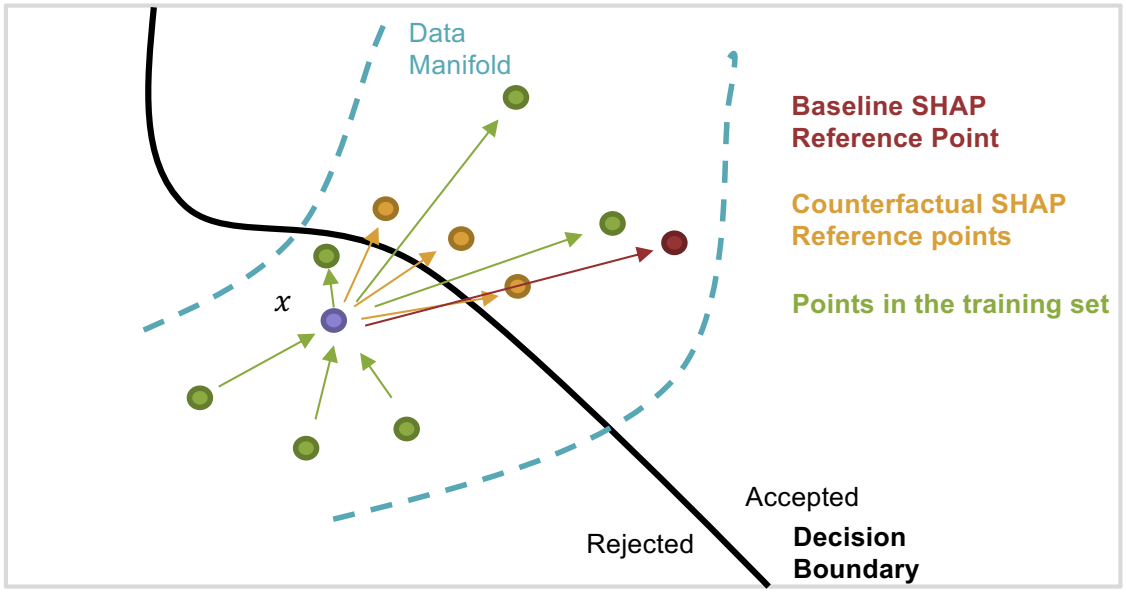
- Target: Risk Management, AI developers
- Questions: *"How does this model work internally?"* or *"What are the features that mostly affect the output of the model?"*

Explanations as **Actionable Recourse**

- Target: Consumers, Regulators
- Questions: *"Why did you reject my loan application?"* or *"What can I do to improve the chances that my application gets approved next time?"*


SHAP Feature Importance and Changes Over Time
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
April
June

Counterfactual SHAP



Baseline SHAP
"which features are important in moving towards the single reference point"

Interventional SHAP
"which features are important in moving up or down the prediction when comparing x with the other samples in the training set"

Counterfactual SHAP
"which features are important in changing the prediction outcome, or in other words in moving on the other side of the decision boundary"

"Counterfactual Shapley Additive Explanations"
Albini Emanuele, Long Jason, Dervovic Danial and Magazzeni Daniele
Workshop on XAI in Finance at ICAIF-2021
https://arxiv.org/abs/2110.14270

Main Challenges with Counterfactual Explanations in Finance

1) Actionability vs "Key Drivers"

Reason codes must be actionable, but they key drivers might be not actionable…

Measuring actionability is also challenging and subjective.

2) Robustness of Explanations

Will the actionable recourse remain so given some changes in the model?

3) Effectiveness of Explanations

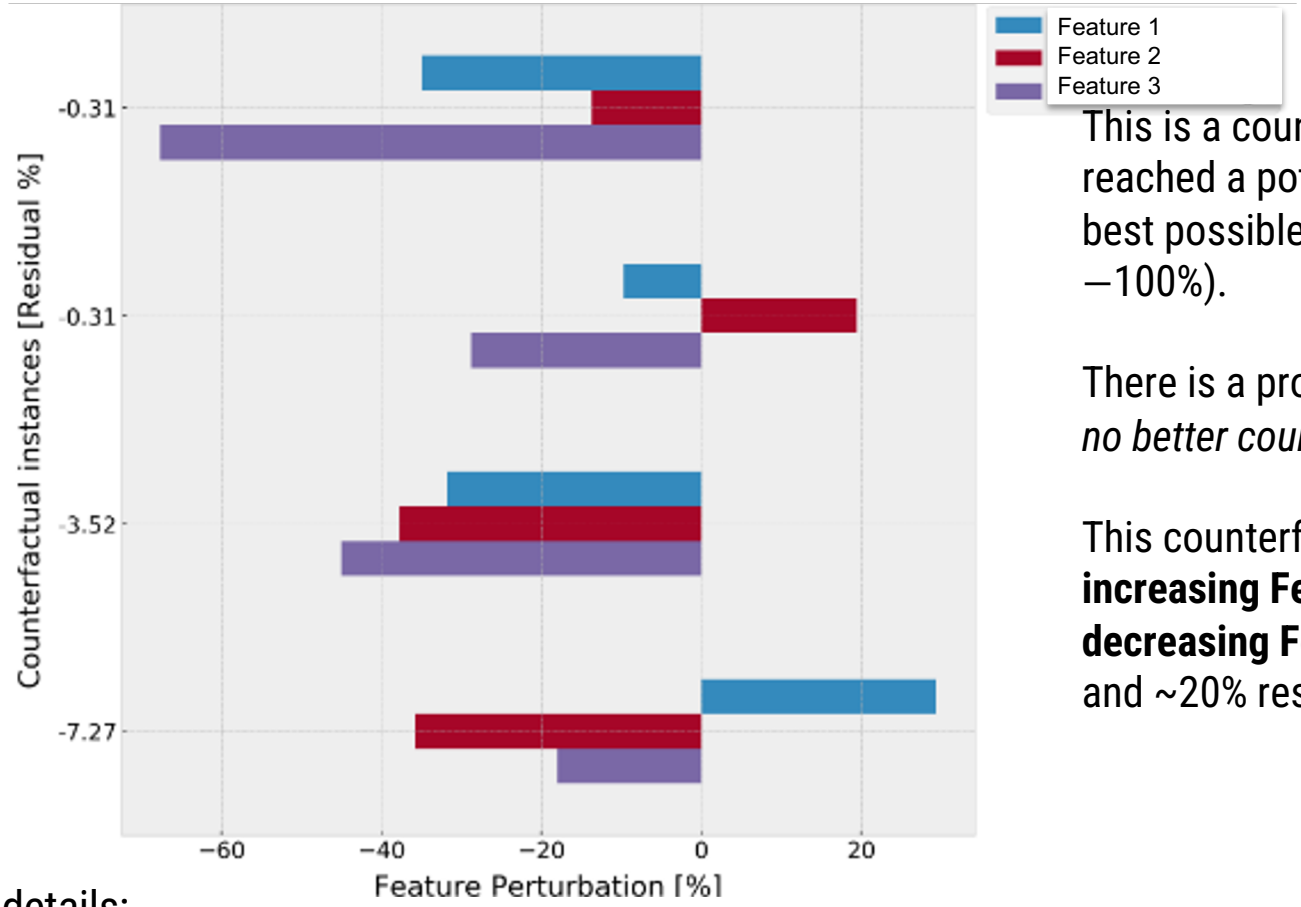
Users must be at the center!

4) Counterfactuals for Regression Models

Huge state space for counterfactuals!

Counterfactuals for Regression Models

- We developed a novel techniques for finding counterfactuals for arbitrary regression models
- Based on Bayesian optimization
- Agnostic about the model



This is a counterfactual that has reached a potential value 7% of the best possible (either +100% or −100%).

There is a probabilistic guarantee that *no better counterfactual exists*.

This counterfactual is achieved by **increasing Feature 1** by ~25% and **decreasing Features 2 and 3** by ~35% and ~20% respectively.

For details:

Counterfactual Explanations for Arbitrary Regression Models

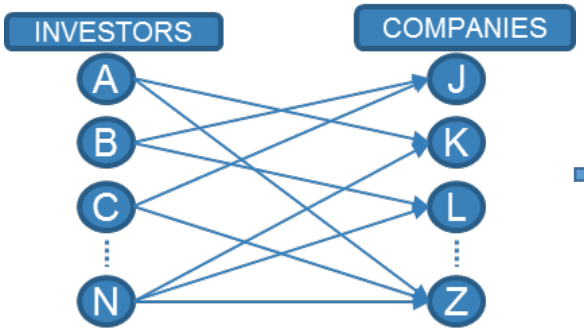
Thomas Spooner, Danial Dervovic, Jason Long, Jon Shepard, Jiahao Chen, Daniele Magazzeni

Explanations for Investor/Company Matching

Goal: build a recommendation system to match companies and investors based on their historical records and preferences

	userId	movieId	rating
0	1	1	4.0
1	1	3	4.0
2	1	6	4.0
3	1	47	5.0
4	1	50	5.0

- Based on the movie/items ratings, recommendations can be made using supervised learning approaches
- Well defined categories for movies/products/services

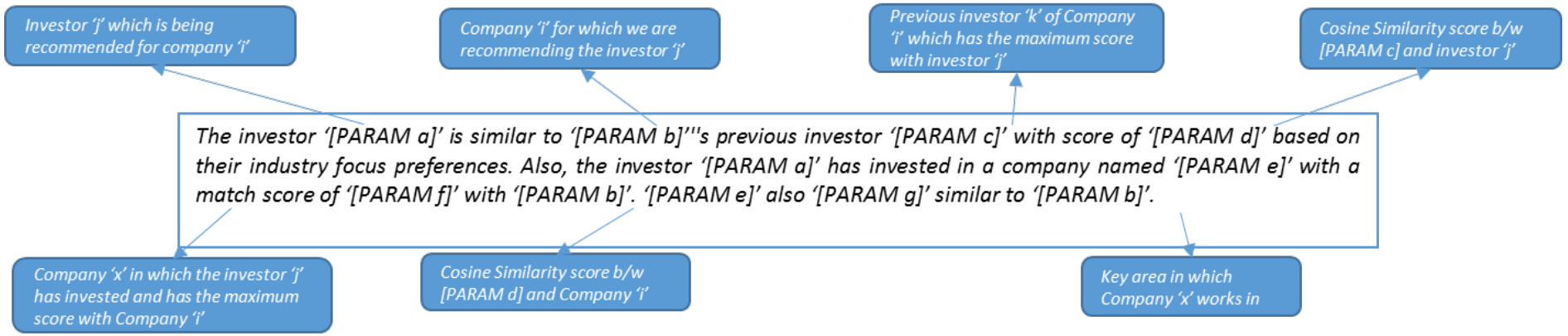


- Only have binary links i.e. if an investor 'i' has invested in company 'j'. No information is there whether an investor likes/dislikes a company and by how much resulting in an UNSUPERVISED RECOMMENDATION MODEL
- How to represent a company/investor with vast categories of each feature such as descriptions/industry focus/locations/funding styles, etc.

Explanations for Investor/Company Matching

Goal: build a recommendation system to match companies and investors based on their historical records and preferences

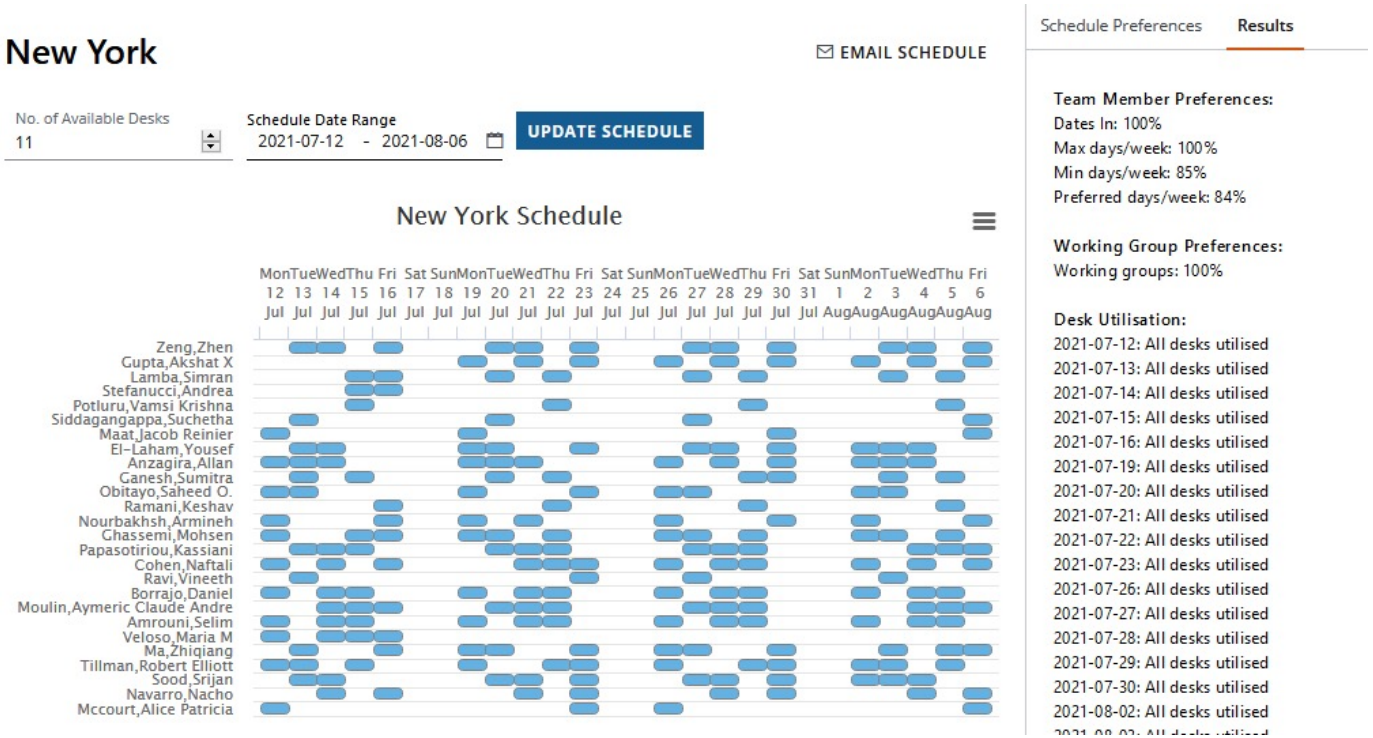
- Provided 105 company-investor matching pairs with scores for business to review
- 64 pairs were confirmed by the business
- 31 pairs were considered not aligned by the business
- Created a parameterized explanatory algorithm to generate explanations for the business as well as clients



Explanations for Resource Allocations

AI for Planning the Return to the Office (RTTO)

Leverage AI to automate, manage constraints efficiently, optimize how teams should rotate in the office



- ✓ Set of hard constraints
- ✓ Set of soft constraints (preferences)
- ✓ Set of meta-constraints (including *fairness*)
- ✓ **Challenge: how to explain the fairness of the proposed schedules**

Synthetic Data for XAI: New Datasets Publicly Available

4 Synthetic DataSets

https://www.jpmorgan.com/technology/artificial-intelligence/initiatives/synthetic-data

- Markets Execution Data
- Payments Data for Fraud Detection
- Customer Journey Events Data
- Anti-Money Laundering (AML) Data

ICAIF'21

2nd ACM International Conference on AI in Finance

ICAIF'21 will be presented virtually November 3-5, 2021,

in association with

The Alan Turing Institute

University of Oxford

Platinum Sponsors

JPMorgan Chase & Co.

usbank

Gold Sponsor

NetApp

Workshop on Explainable AI in Finance

@ ICAIF 2021

Workshop Co-Chairs: Anupam Datta (Carnegie Mellon University), Himabindu Lakkaraju (Harvard University), Daniele Magazzeni (J.P. Morgan AI Research), Francesca Toni (Imperial College London)

9:05 am - 10.05 am
Panel Discussion

Explainable AI from Regulatory and Risk Management Perspective

Albert Chang (Senior Counsel at CFPB)
Melissa Koide (CEO & Director at FinRegLab)
Walter McCahon (Principal: Data Protection and Ethics at UK Finance)
David Heike (J.P. Morgan Risk Modeling)




10:15 am - 11:30 am
Session 1

Long Presentations (methods)

10:15 am - 10:30 am Counterfactual Shapley Additive Explanations
10:30 am - 10:45 am Leveraging Causal Relations to Provide Counterfactual Explanations and Feasible Recommendation to End Users
10:45 am - 11:00 am A Survey on the Robustness of Feature Importance and Counterfactual Explanations
11:00 am - 11:15 am Causal Injection into Neural Networks
11:15 am - 11:30 am Principled Diverse Counterfactuals in Multi-Linear Models

11:40 am - 12:25
Session 2

Long Presentations (case studies)

11:40 am - 11:55 am Explanations in Whose Interests?
11:55 am - 12:10 pm An Investigation of the Impact of COVID-19 Non-Pharmaceuticals Interventions and Economic Support Policies on Foreign Exchange Markets with Explainable AI Techniques
12:10 pm - 12:25 pm Parameterized Explanations for Investor / Company Matching

12:30 pm - 1:00 pm
Session 3

Short Presentations (methods and case studies)

12:30 pm - 12:35 pm Designing Inherently Interpretable Machine Learning Models
12:35 pm - 12:40 pm Neural Policy Learning in Interpretable Trading Strategies Using Inductive Priors
12:40 pm - 12:45 pm Fairness Through Time
12:45 pm - 12:50 pm Logical English for Legal Applications
12:50 pm - 12:55 pm Venture Capital (Mis)Allocation in the Age of AI
12:55 pm - 1:00 pm Building Effective Guidelines for XAI: Lessons from Philosophy

1:00 pm
Concluding remarks

Closing Remarks

Daniele Magazzeni (AI Research Director and Head of Explainable AI Center of Excellence at J.P. Morgan) &
Francesca Toni (Imperial College London)

Opportunities at J.P. Morgan AI Research

Bank Street, Canary Wharf, London

- Internships
- PhD Fellowships
- Full-time positions

- In London and New York

Thank you!

Daniele Magazzeni, PhD
AI Research Director
Head of Explainable AI Center of Excellence
J.P. Morgan

daniele.magazzeni@jpmorgan.com