Explainability in Graph-based Machine Learning and Social Network Analysis
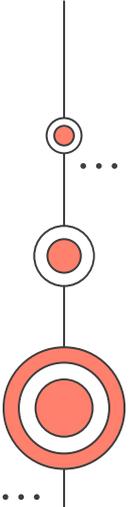
Sophie Sadler
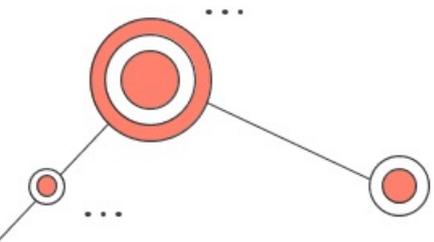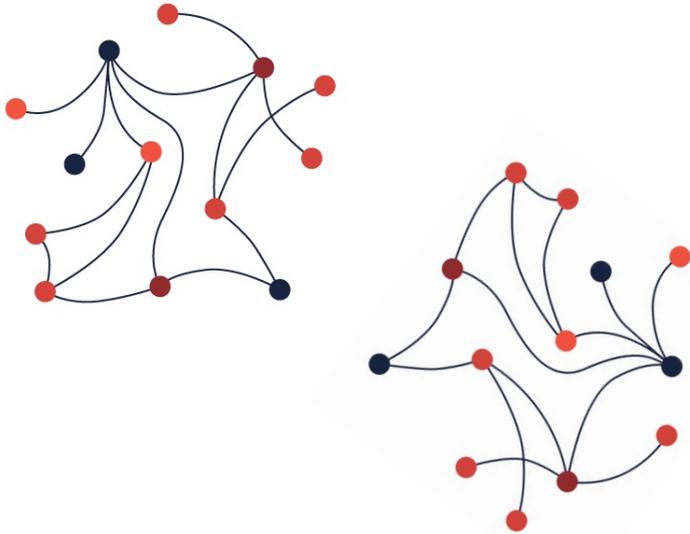
Agenda
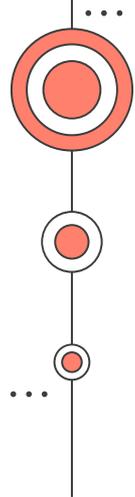
01
Introduction

Who I Am

Currently:
- 3rd year Computer Science PhD @ Swansea University
- Research Collaborator @ Meta (previously Facebook)

Previously:
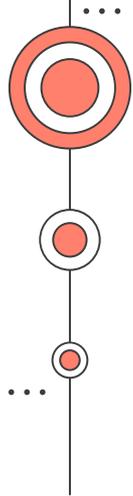- Undergraduate Mathematics @ Uni of Cambridge
- AI Resident @ Microsoft Research

Graphs and Network Data

- Relational data
- Data points known as "nodes"
- Connections known as "edges"
- SOMETIMES: additional information known as "attributes"

Examples:
- Social networks
- Chemistry – interaction pathways, compounds
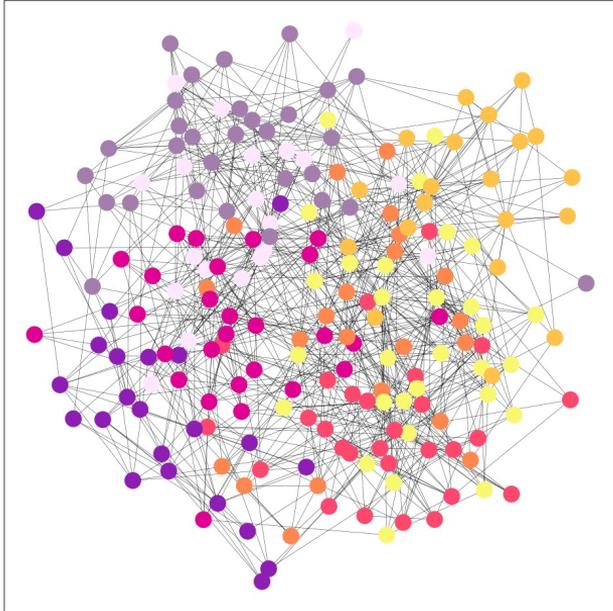- Public health
- Computing

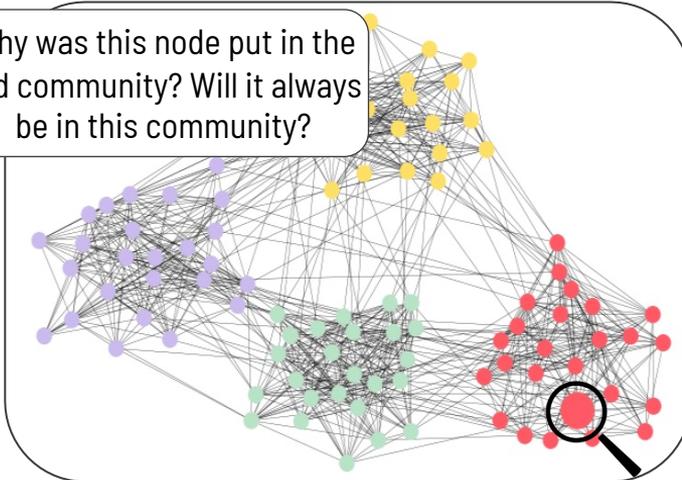# 02
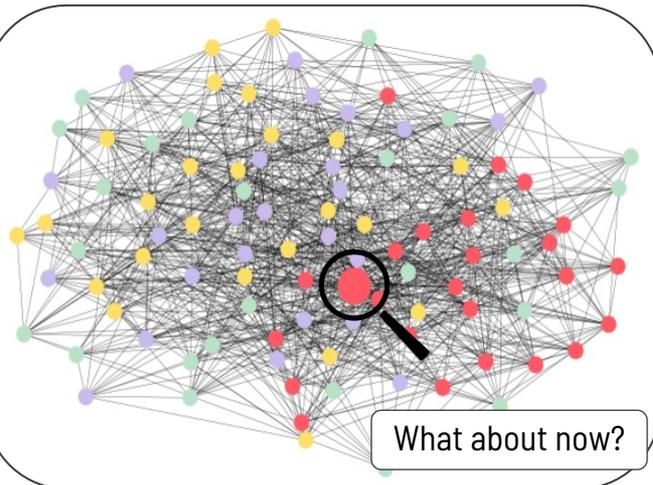## Explainable Community Finding

Community Finding

- Communities: areas of **high connectivity**

- E.g. research fields in a network representing collaborations!

- Algorithms are normally **complex**, **stochastic** and **lacking in explanation**

Explainable Community Finding

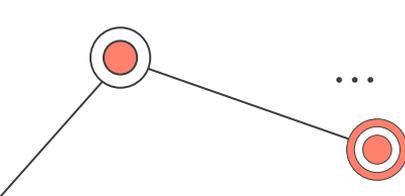**Similarly:**
- Why are these two nodes in the same/different communities?
- Why does this set of nodes form a community?

Explainable Community Finding

Aim: to generate explanations that are simpler to understand than the algorithm.
Idea: identify easily understood features that can explain the algorithm outputs

Longlist of features (e.g):

Cut ratio

Relative betweenness
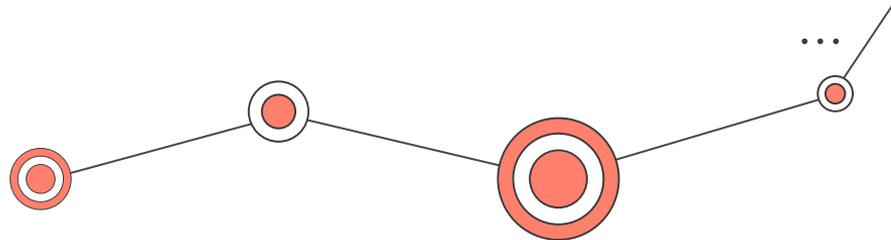
Relative diameter

Relative degree

Our model-agnostic
methodology
→

Explainable features (e.g.):
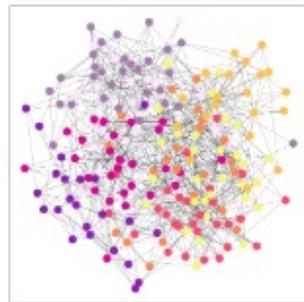
Relative betweenness

Cut ratio

Methodology

Explainable features on three levels: node, node-pair, community

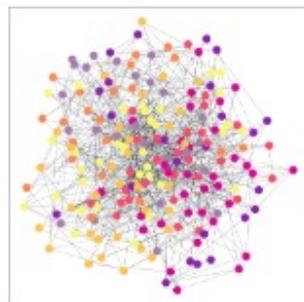Train a random forest for three binary classification problems:
1. Is this node normally in the same community every time or not?
2. Do these two nodes belong to the same community or not?
3. Is this set of nodes a community or not?

Start with a longlist of easily interpretable features
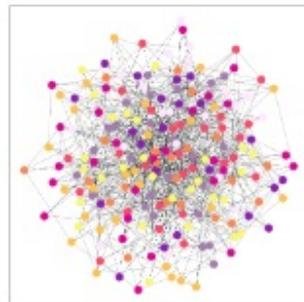
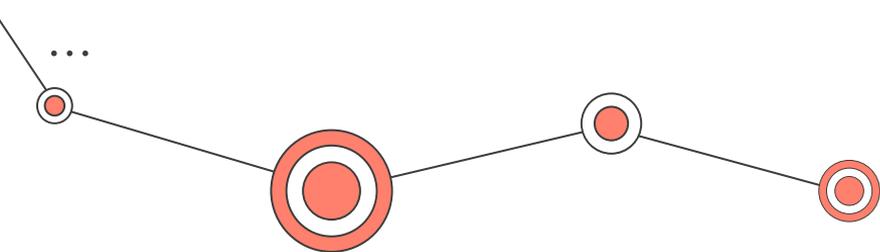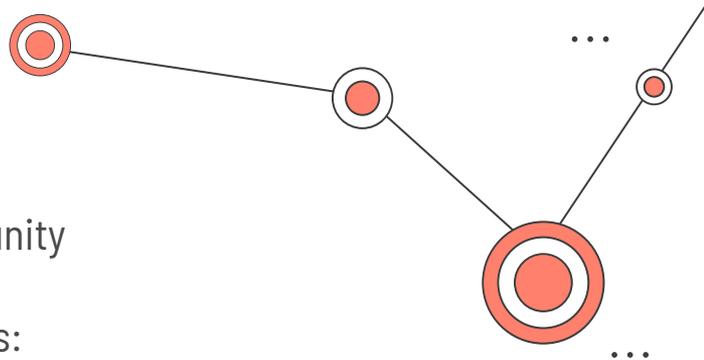Use permutation importance to determine which are explainable

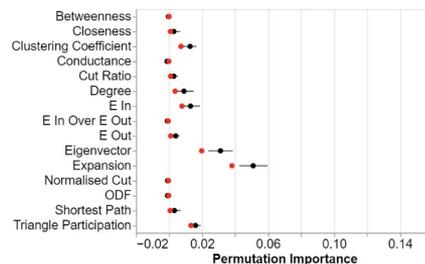(a) $\mu = 0.2$ (b) $\mu = 0.3$ (c) $\mu = 0.4$

Analysis

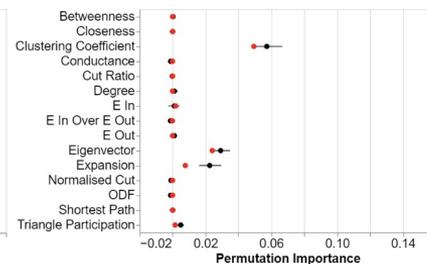Node and node-pair level: surprisingly algorithm-independent
Paper in preparation

Community level: Cut ratio; internal-external; for high μ, relative betweenness
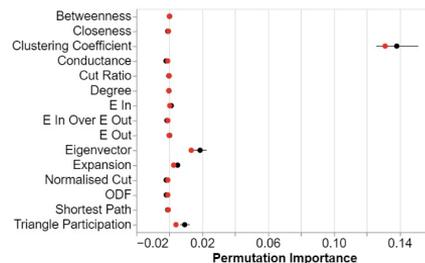**Paper accepted to Complex Networks 2021**
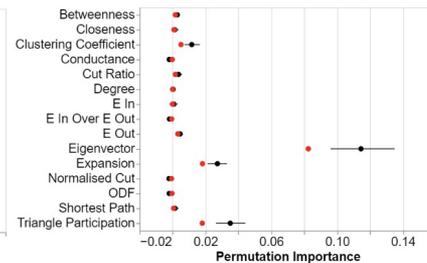**Work presented in a poster session @ GEM ECML Workshop**

(a) Infomap, $\mu = 0.2$

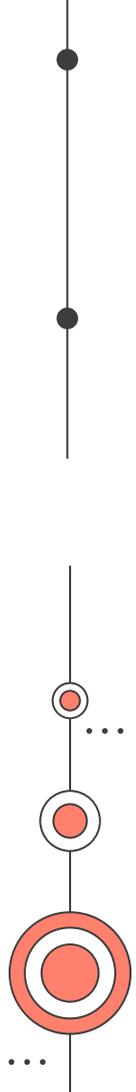(b) Infomap, $\mu = 0.3$

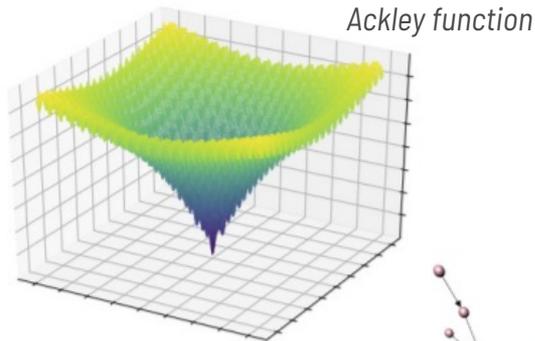(c) Infomap, $\mu = 0.4$

(d) Louvain, $\mu = 0.2$

03

Extrema Graphs

What is a LON?

Ackley function

Ackley function represented as a LON

Optimization algorithms stochastically find minimal solutions from a fitness landscape. Visualization required.

LON: Local Optima Network

Nodes represent maxima or minima

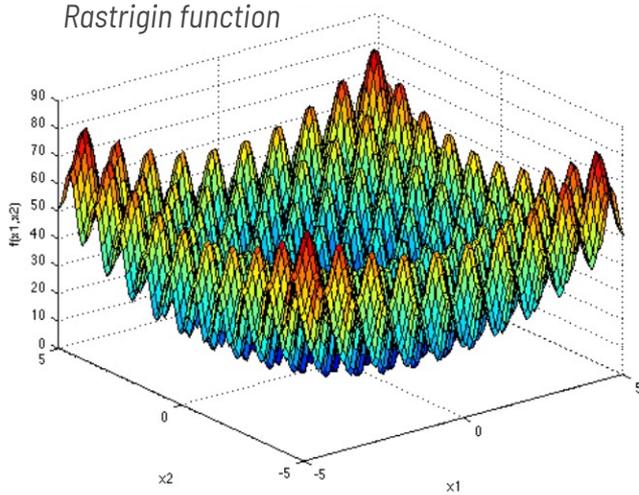Edges represent how "easy" it is to transition from one optimum to another
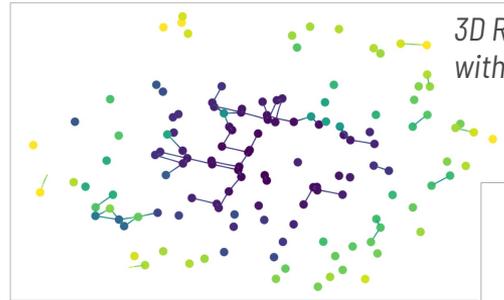
Extrema Graphs

Our contribution: extrema graphs

Differences from LONs:
- Both maxima and minima
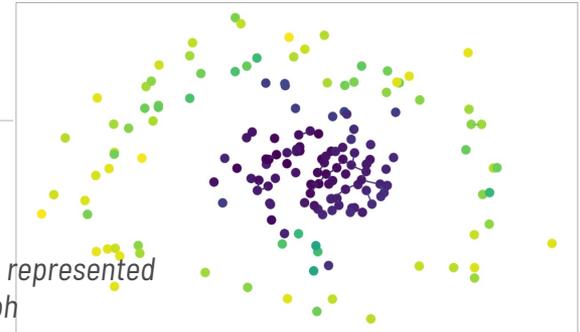- Project to 2D using MDS - generalise to many dimensions

Work in Progress

Rastrigin function

$$f(\mathbf{x}) = 10d + \sum_{i=1}^{d} \left[x_i^2 - 10\cos(2\pi x_i) \right]$$

3D Rastrigin function represented with an extrema graph

5D Rastrigin function represented with an extrema graph
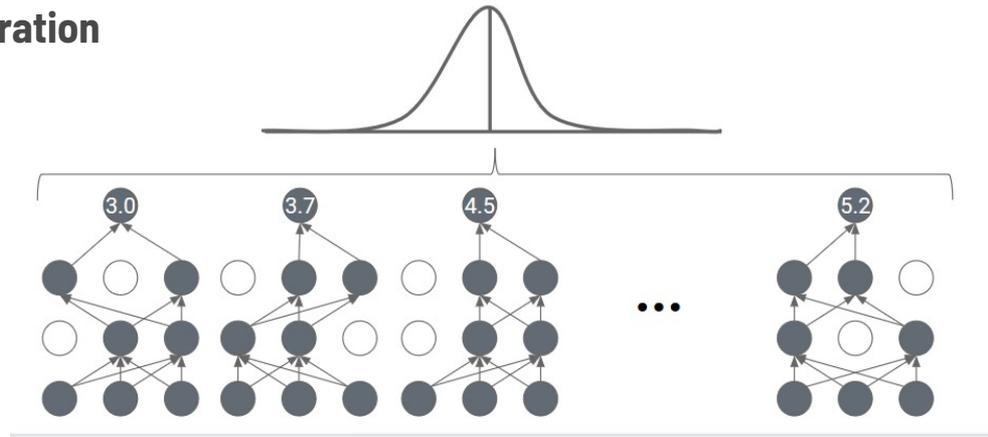
04

Representing Uncertainty in Social Network Virality Predictions

Modelling Virality

- Working in collaboration with Meta (formerly Facebook)
- Existing models predict how many views a piece of content will receive
- Project aim: model uncertainty in these predictions
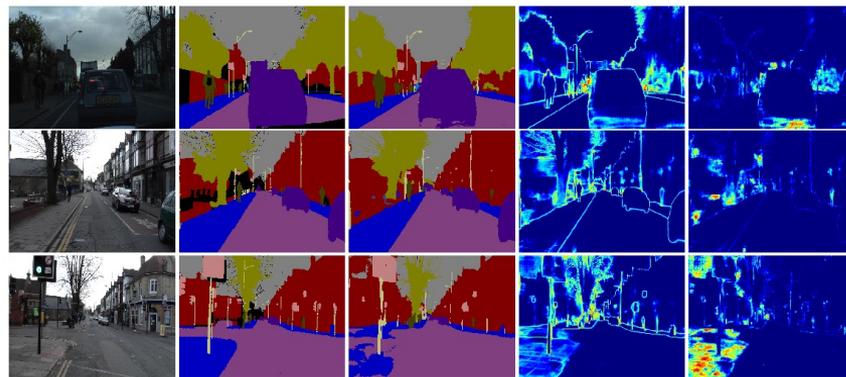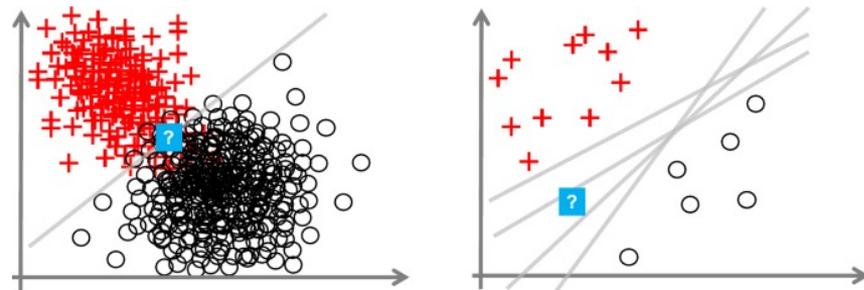
Paper in preparation

Types of Uncertainty

Epistemic Uncertainty:
- Captures the imperfection of the model
- This is what MC Dropout models
- If we show the model more data, this should decrease

Aleatoric Uncertainty:
- Captures noise due to lack of information
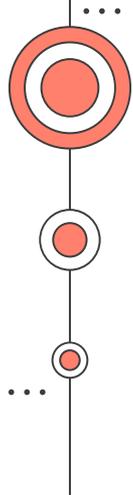- Showing the model more data will make no difference

(a) Input Image   (b) Ground Truth   (c) Semantic Segmentation   (d) Aleatoric Uncertainty   (e) Epistemic Uncertainty
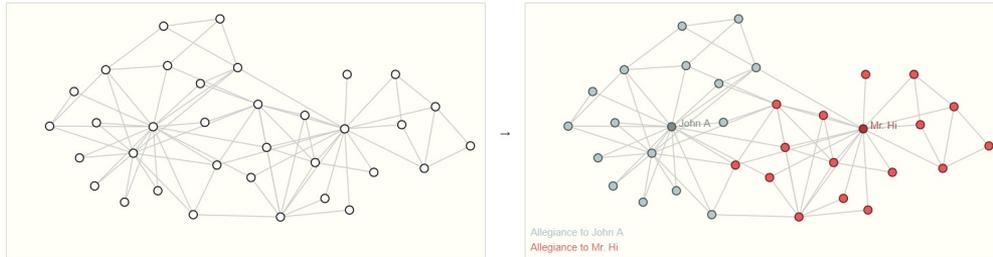
05

Graph Neural Networks

Prediction Tasks on Graphs

As with images, we can perform prediction tasks on graphs

- Which type of chemical compound does this graph represent?
- Which class does a certain node belong to?
- Which class does a certain edge belong to?

... analogous to:

- Cat vs dog images
- Semantic segmentation
- ?

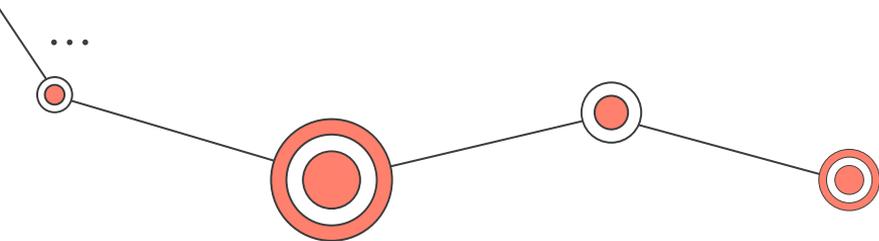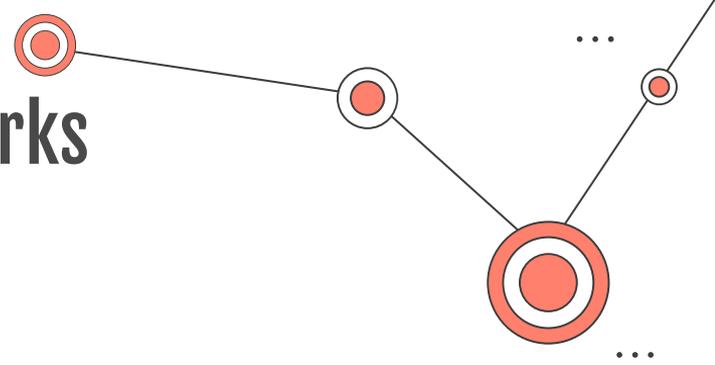Allegiance to John A
Allegiance to Mr. Hi
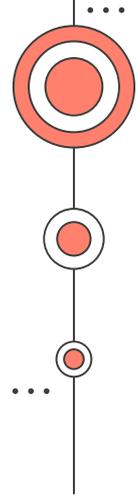
Graph Neural Networks

Normally use CNNs for image prediction tasks.

For graph prediction tasks, use analogous GNNs.

Graph data is complex, as we need to represent nodes, edges, node-edge connections, and attributes
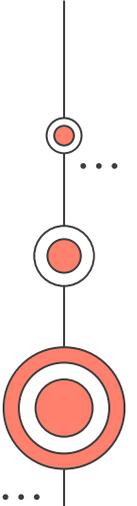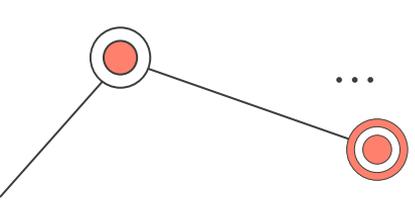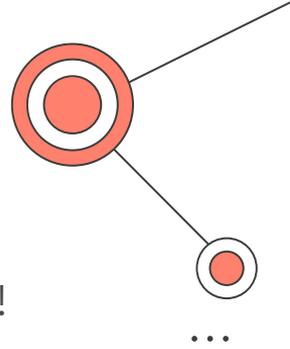
Future work

06
Conclusion

Conclusion

Many areas for incorporating explainability into graph-based ML or social network analysis!

- Explainable community finding
- Extrema graphs for understanding fitness landscapes
- Modelling and understanding uncertainty in predictions of content virality
- Explainable graph neural networks and multivariate analysis
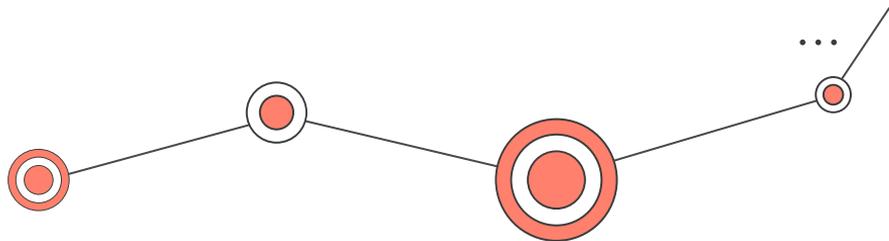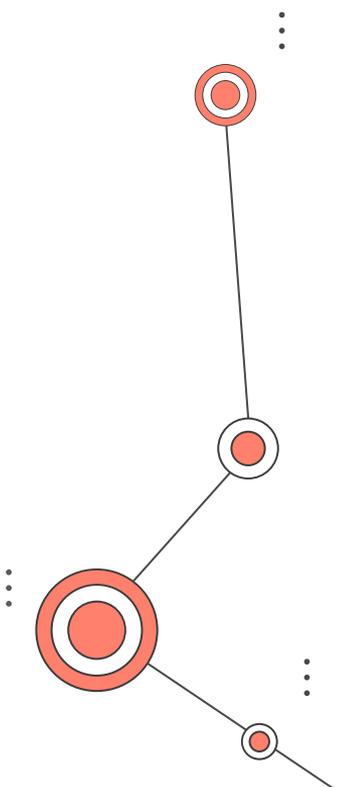
Thanks!

Thanks to: Daniel Archambault, Derek Greene, Alma Rahat, David Walker, Dima Karamshuk, Ami Tavory, Milan Vojnovic

Questions?

sophiesadler@fb.com
1915596@swansea.ac.uk