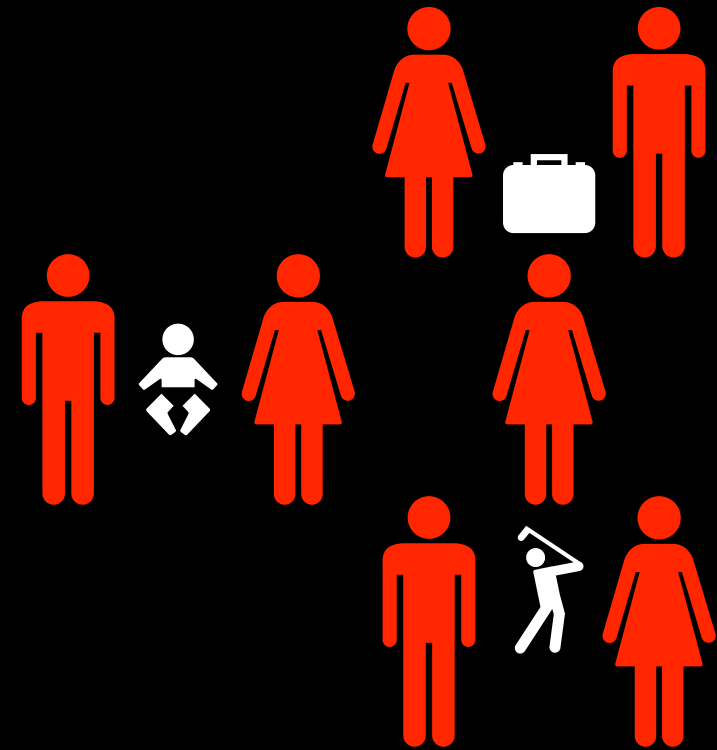


Learning the Future Spread of an Epidemic

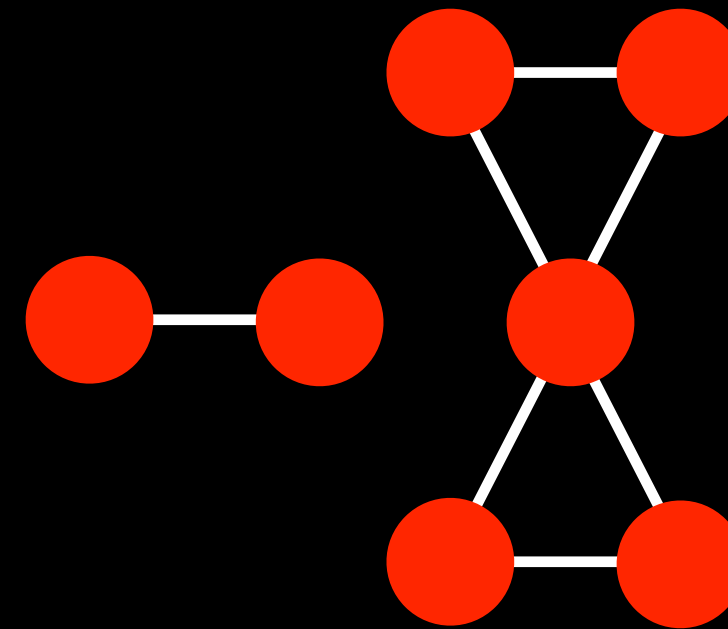
Supervised by Professor Simon Dobson

Simon Buchacher, 25 August 2021

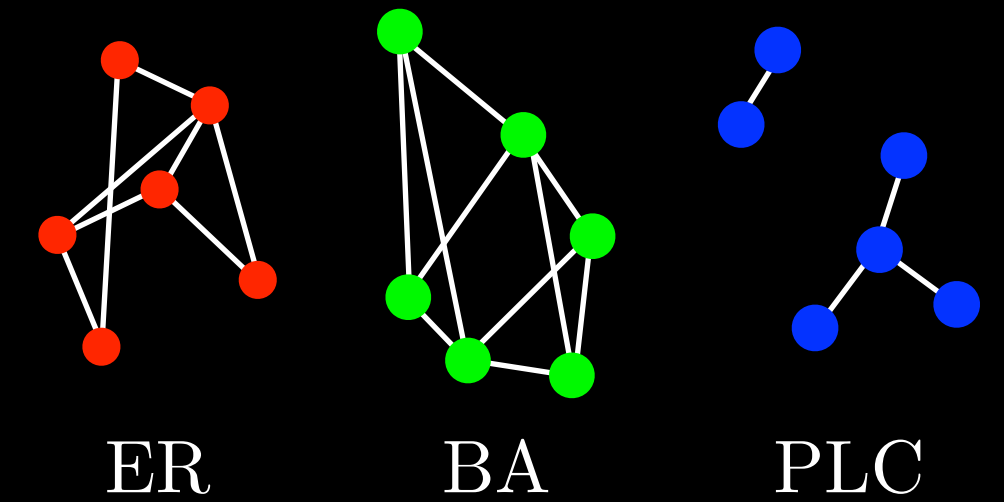
Simulation



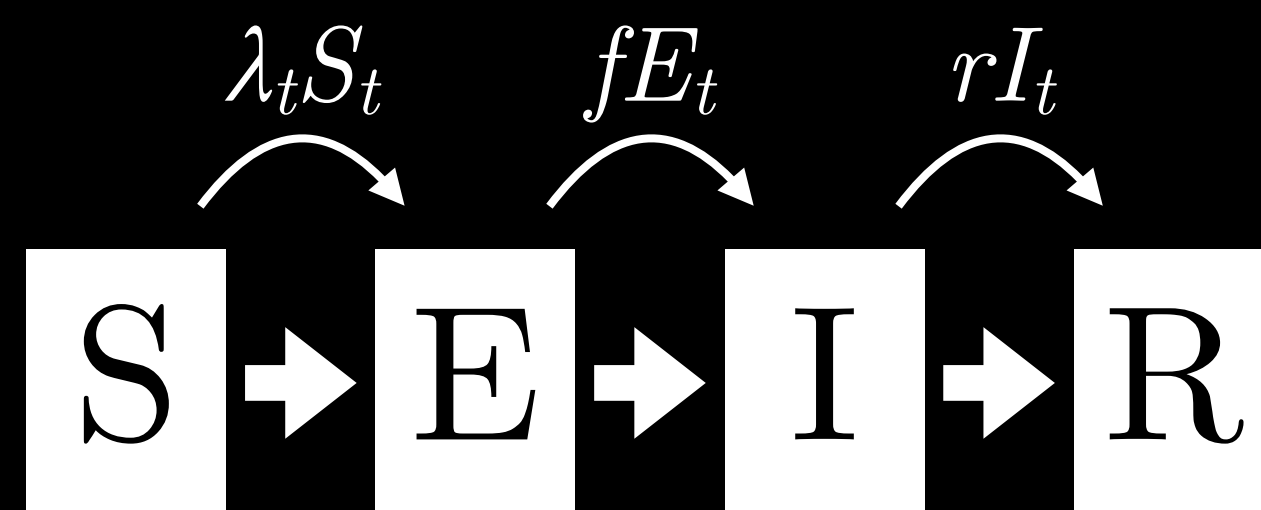
Population



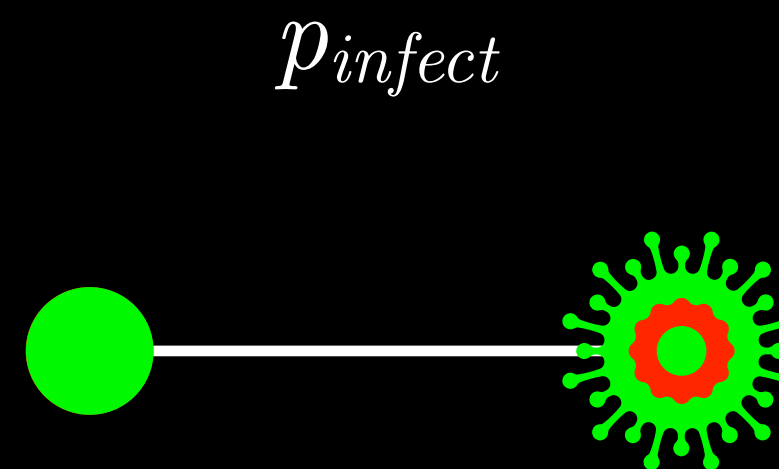
Human Contact Network



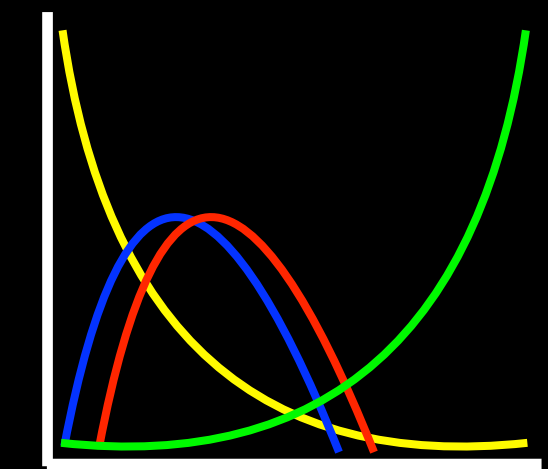
Network Types



Compartmental Model

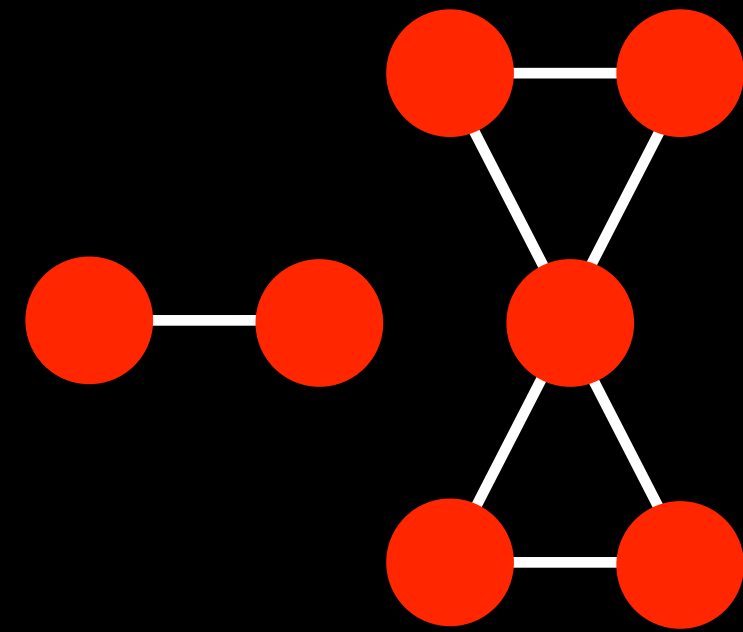


Network Case

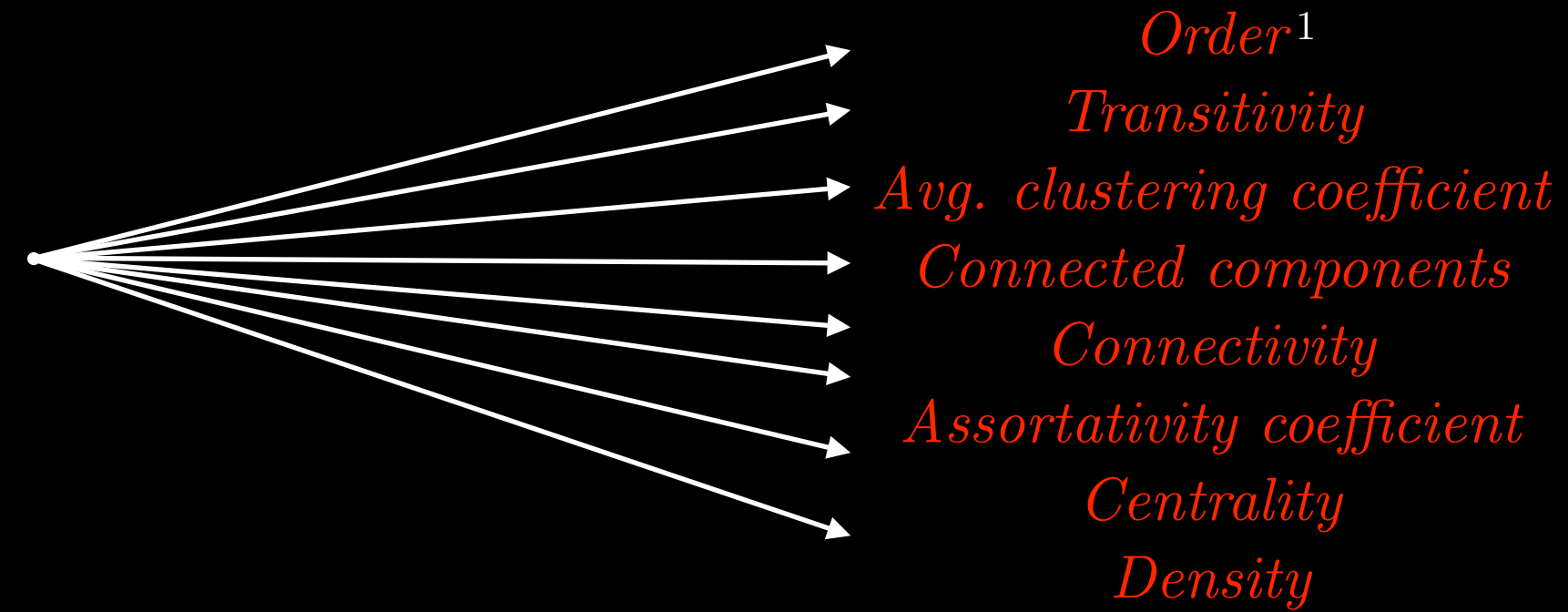


Simulated Epidemic

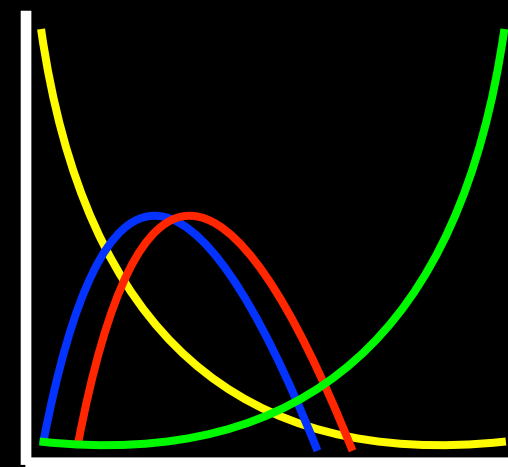
Dataset



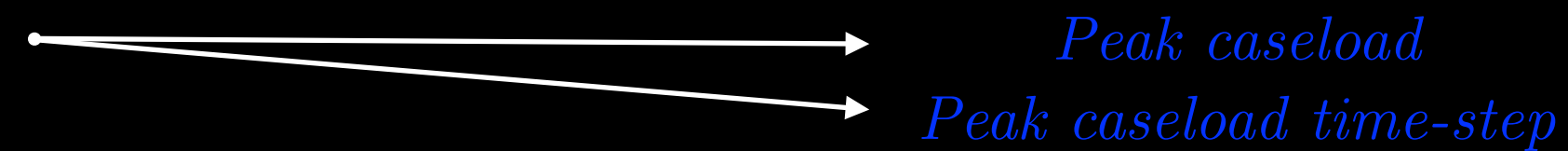
Human Contact Network



Network Topological Features



Simulated Epidemic



Targets

¹To be exact, order — or, the number of nodes in the network — is not extracted from the network but is a parameter passed to its generator.

Dataset



University of
St Andrews

Machine Learning

8802	0.008	1	0.208	TRUE	0.002	9.010	0.008	3444	599
6334	0.011	1	0.491	TRUE	-0.004	1.689	0.011	2839	596
9020	0.003	1	0.393	TRUE	-0.003	2.999	0.003	4333	646
5535	0.015	1	0.394	TRUE	0.003	2.649	0.014	2052	545
...

Dataset

8802	0.008	1	0.208	TRUE	0.002	9.010	0.008	3444	599
6334	0.011	1	0.491	TRUE	-0.004	1.689	0.011	2839	596
9020	0.003	1	0.393	TRUE	-0.003	2.999	0.003	4333	646
5535	0.015	1	0.394	TRUE	0.003	2.649	0.014	2052	545
...

Input and Output

$$Y = f(X, \theta) + \epsilon$$

General Prediction Equation

$$J(\theta) = \frac{1}{2} \sum_{i=1}^n (\underbrace{h_{\theta}(x^{(i)})}_{\text{blue}} - \underbrace{y^{(i)}}_{\text{red}})^2$$

$\epsilon = Y - f(X, \theta)$

Cost Function Optimisation

Test MSE	R^2
4921.13	-0.04

Model Evaluation



Feature Importance

Learning the Future Spread of an Epidemic

Supervised by Professor Simon Dobson



University
of
St Andrews

Simon Buchacher

August 2021

MSc in Artificial Intelligence

Abstract

The current pandemic has made clear how little we know about the detailed spread of diseases in populations. In this project, the spread of infectious diseases through closed populations is modelled by generalising the Susceptible-Exposed-Infectious-Removed (SEIR) compartmental model to the individual-based network case. Large-scale epidemic spreading simulations are performed over three types of human contact networks: Erdős-Rényi networks, Barabási-Albert networks and power-law with cut-off networks. Persons in the population are represented as nodes, interpersonal relationships as edges in the network. During simulation, individuals' symptom and infection statuses as well as their transitions from one compartment to the next are tracked collectively. In addition, network topological features of the underlying human contact network are extracted. Traditional machine learning techniques applied to epidemic spreading over networks are explored, evaluated and utilised to predict two targets, the peak caseload and the time-step at which the peak caseload is reached, as well as to analyse the importance of the extracted network topological features for target prediction.