# WEBINAR STARTS AT 14:00

agilysis

# ARTIFICIAL INTELLIGENCE FOR COMPLEX ANALYSES

CRAIG SMITH & BRUCE WALTON

# WEBINAR SUPPORT

- Please use the Q&A Section to ask questions – we will answer as many as we can
- This is being recorded and will be available to review shortly
- The PDF slides are also available

# ARTIFICIAL INTELLIGENCE

- Covers a lot of different techniques and algorithms
- We can’t cover all of them in a 30-minute webinar
- So we will focus on a flexible framework that serves as the basis of many AI algorithms and techniques: ***artificial neural networks***

agilysis

# INTRODUCTION

- What are neural networks and why do they work?
- What sort of things can we use them for?
- Where do I start if I want to do this?

# WHAT IS A NEURAL NETWORK?

Pictures of a fixed size

Does the picture contain a dog?

Addition

Yes/No

$\{0,1\}$

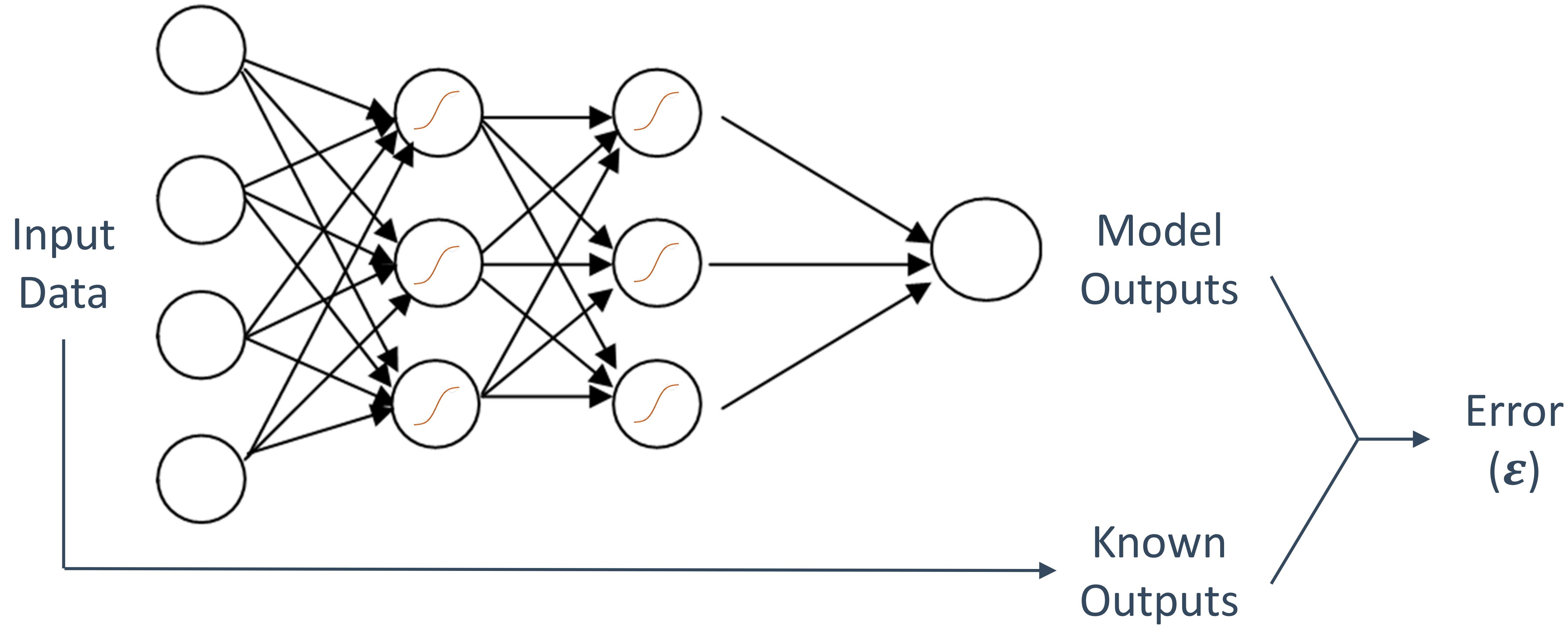
$f$

$\{x_1, x_2, x_3, \dots, x_n\}$

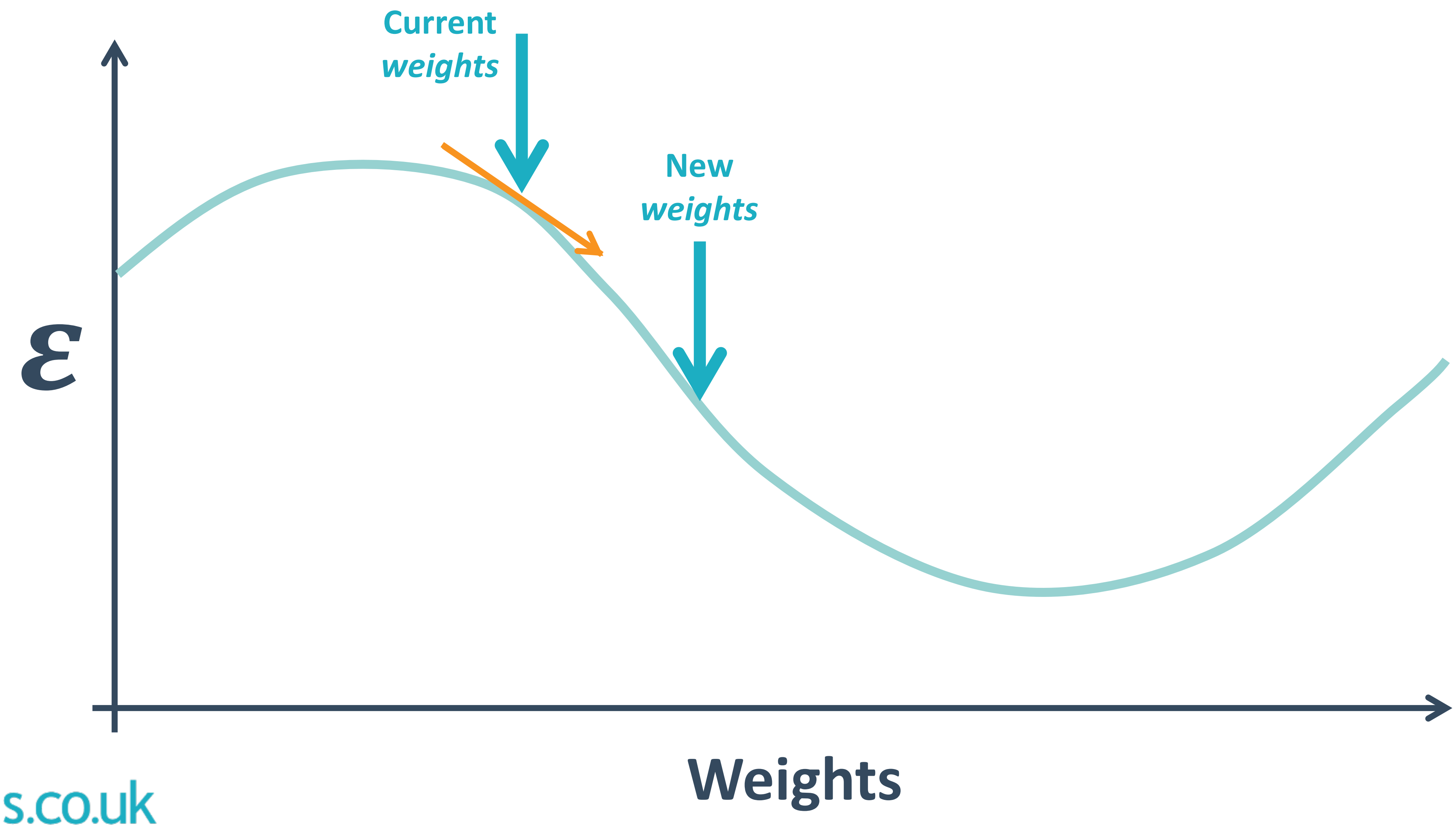
$\{y_1 = F(x_1),$
$y_2 = F(x_2),$
$y_3 = F(x_3), \dots,$
$y_n = F(x_n)\}$

**Goal:** Minimise error ($\boldsymbol{\varepsilon}$) between $f$ and $F$ on known data

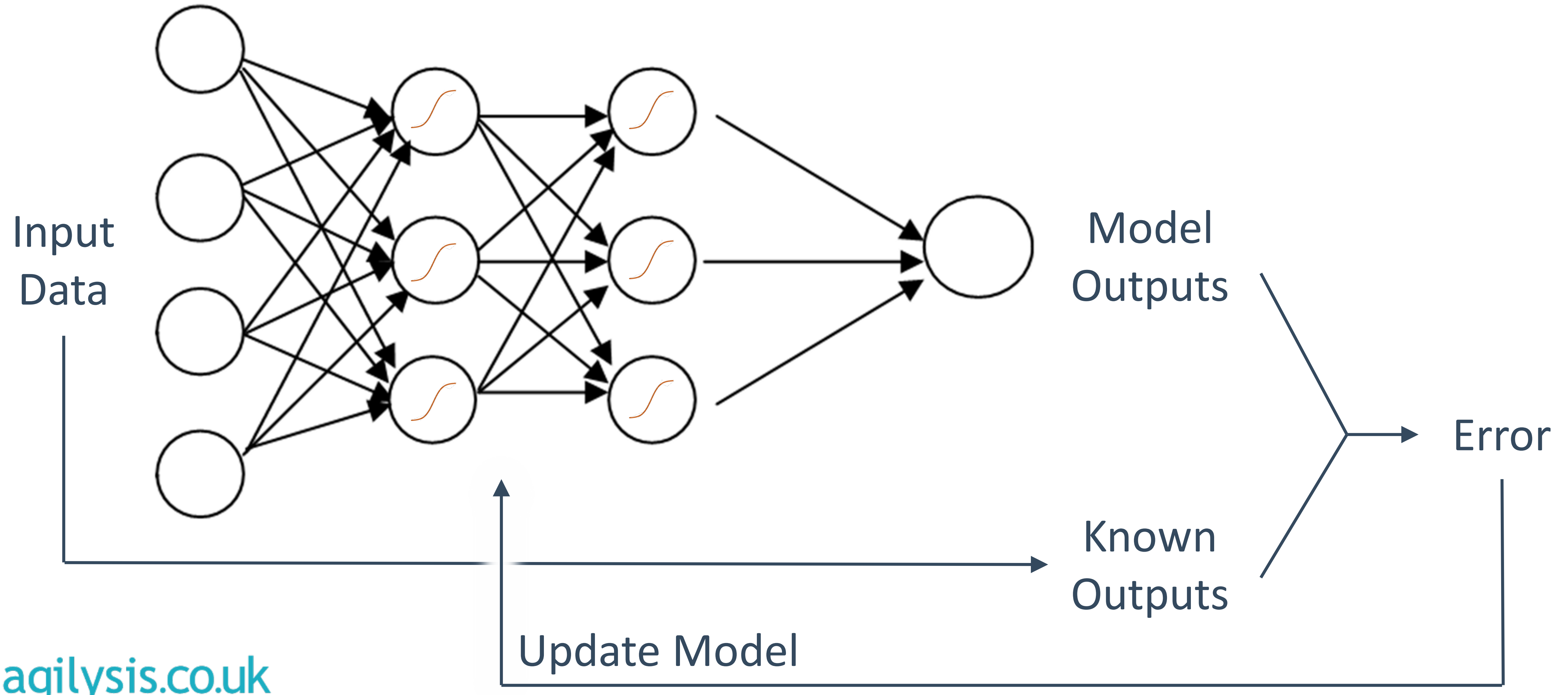
# WHAT IS A NEURAL NETWORK?



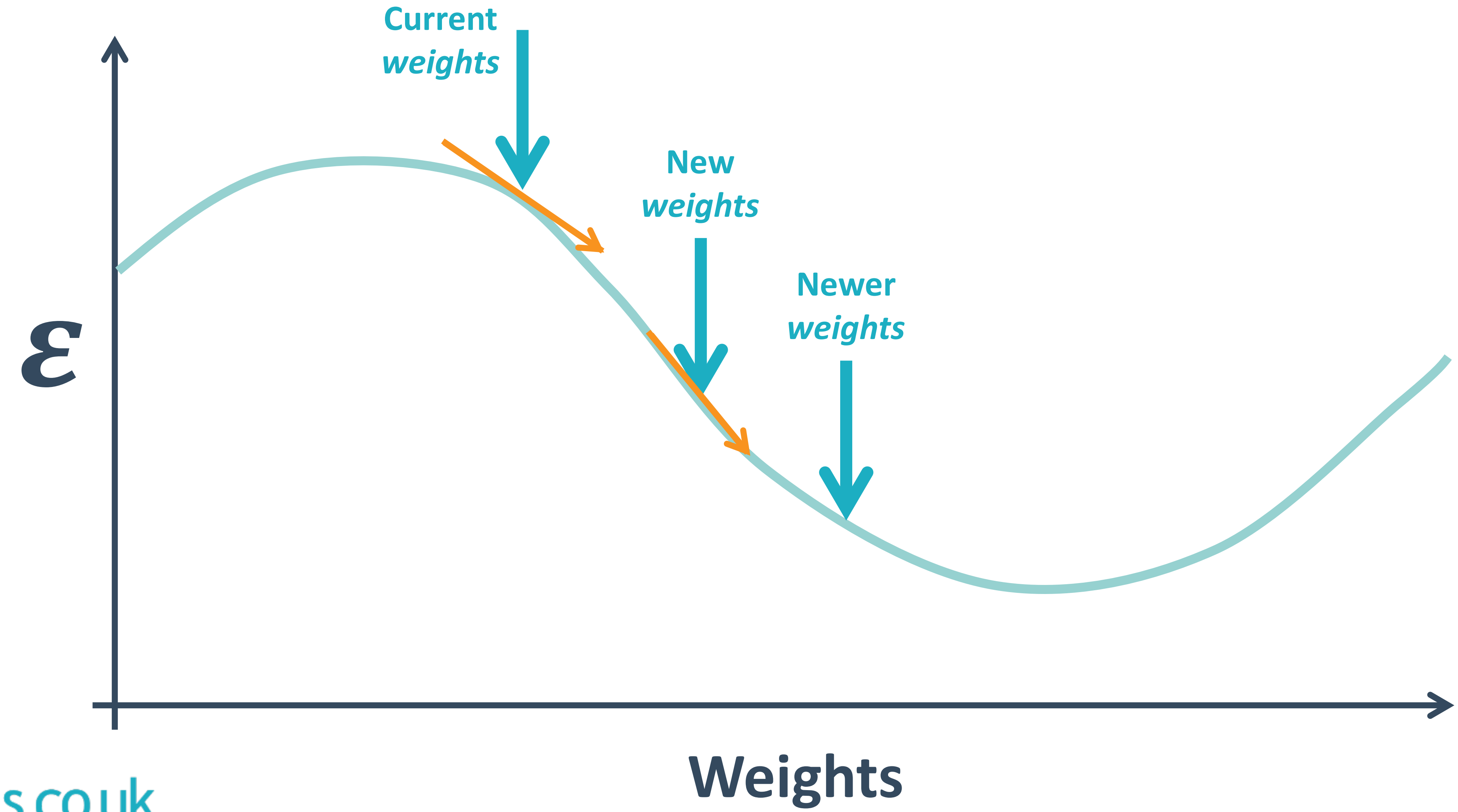
# WHAT IS A NEURAL NETWORK?



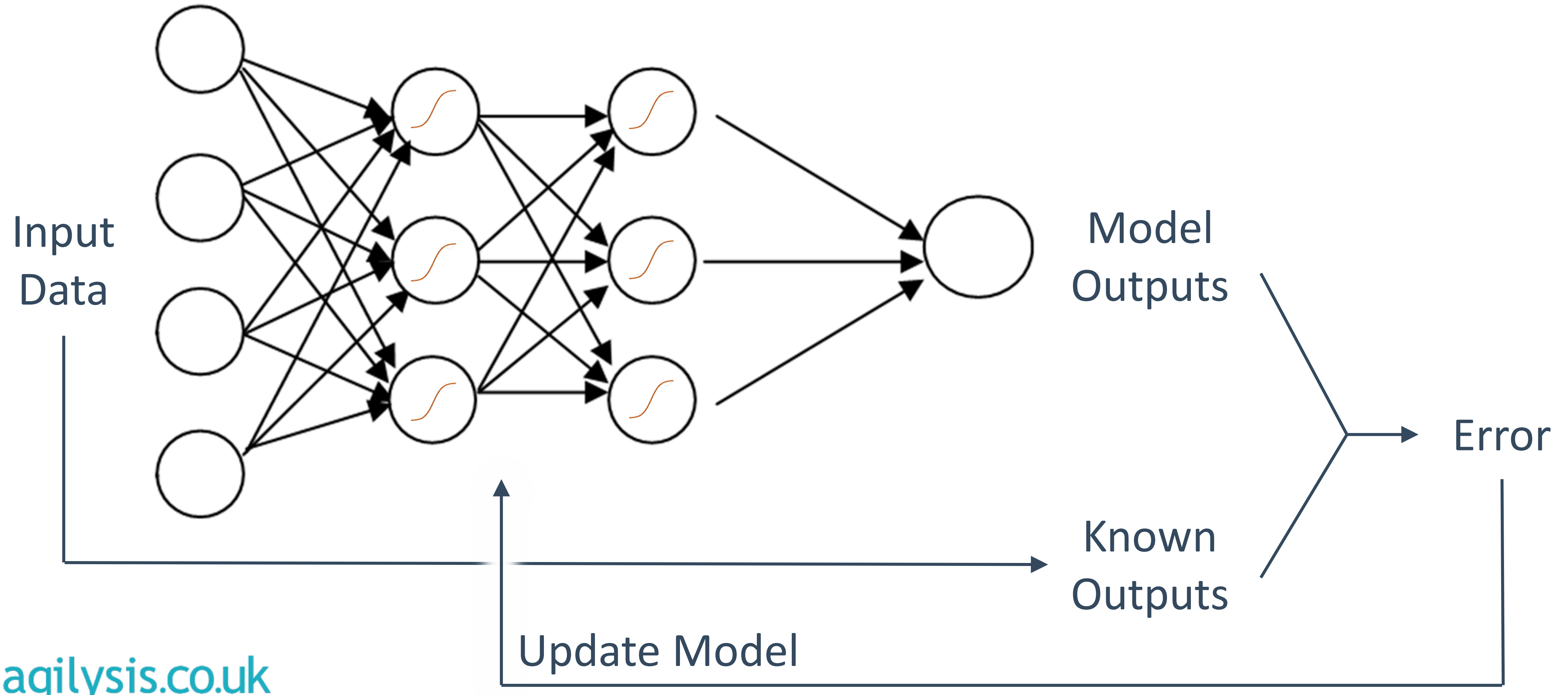
# WHAT IS A NEURAL NETWORK?



# WHAT IS A NEURAL NETWORK?

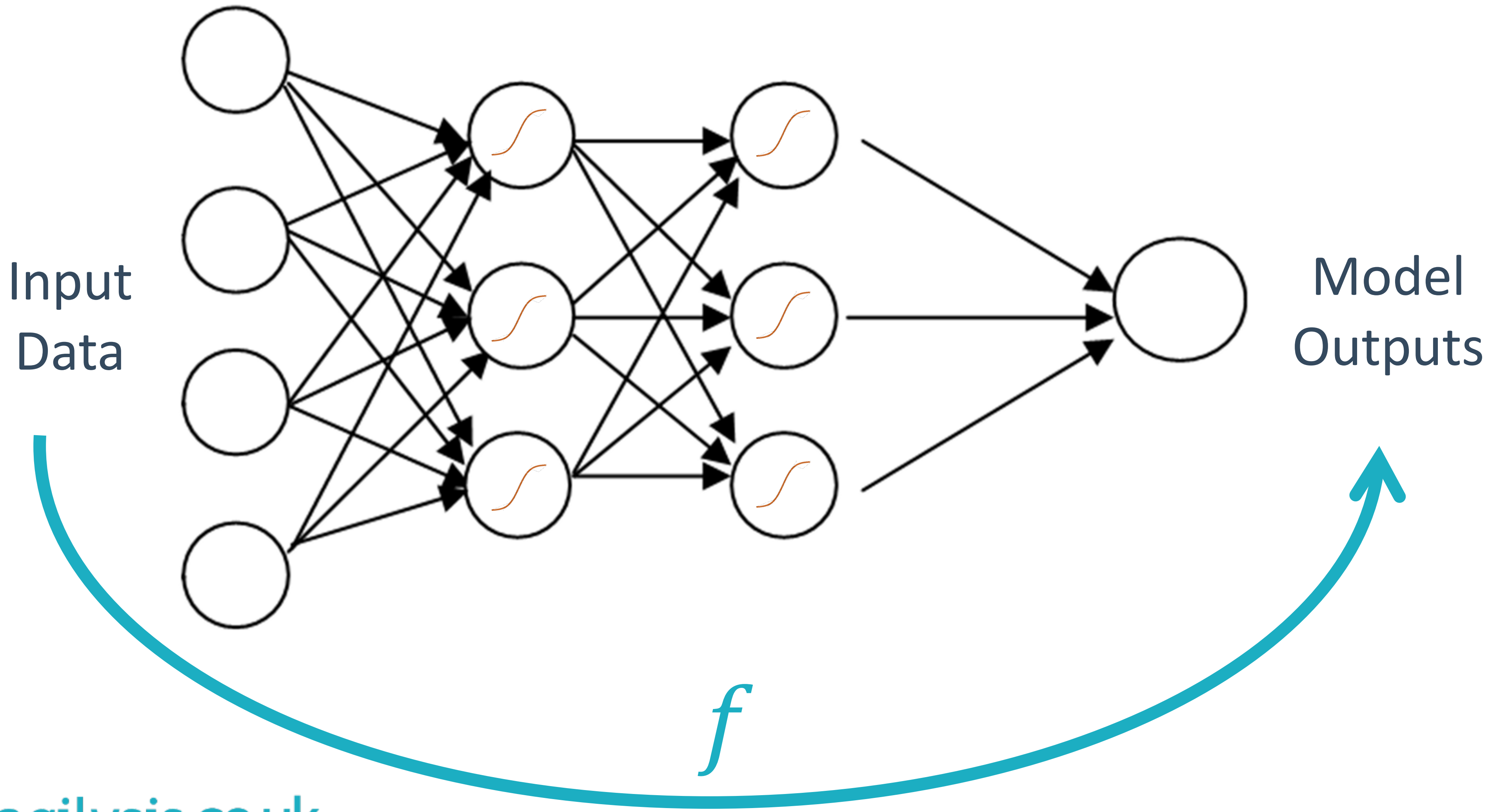


# WHAT IS A NEURAL NETWORK?

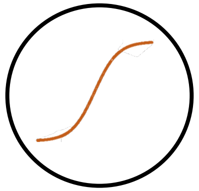


# WHAT IS A NEURAL NETWORK?

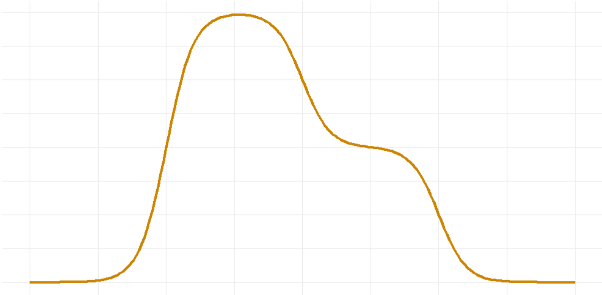
agilysis

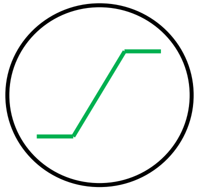


# WHAT IS A NEURAL NETWORK?

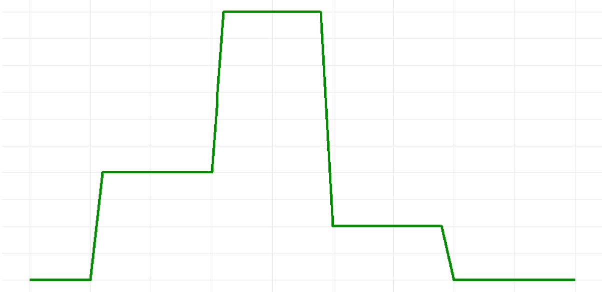


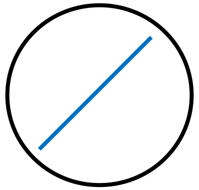
- *Sigmoid* activation builds the model out of curvy pieces



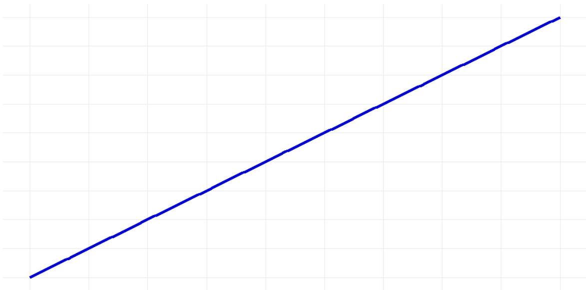


- *Relu* activation builds the model out of piecewise-linear pieces





- *Linear* activation builds the model out of linear pieces *(hence is linear)*

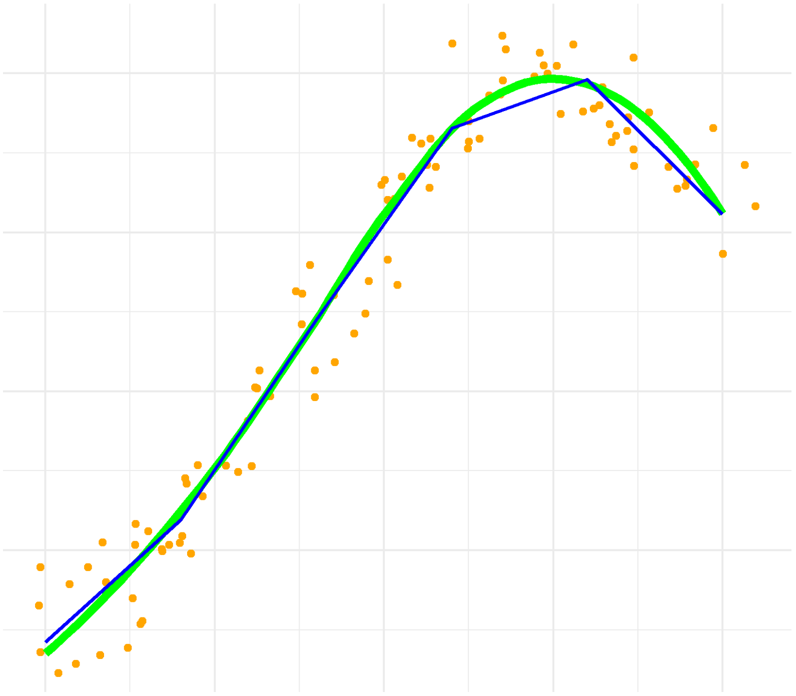


# WHAT CAN WE USE THEM FOR?

- Regression models

# (NON-LINEAR) REGRESSION

- Use a combination of discrete and continuous input variables
- Predicting continuous output values
- Interpolate data to cover blackspots
- Can have issues with extrapolation



# TFL ROAD DANGER REDUCTION

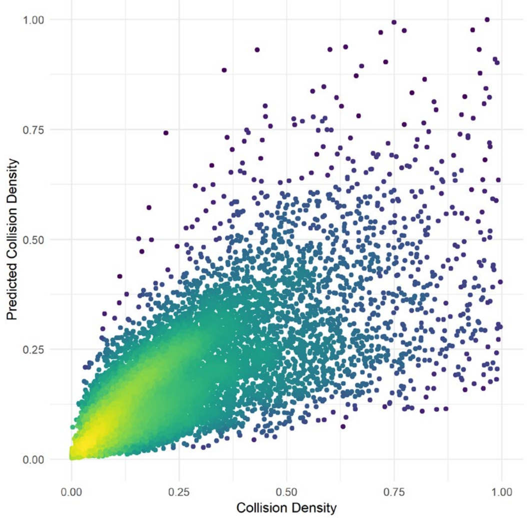
- Use historic STATS19 collision data, matched to a granular road network to calculate measures of risk:
  - *Collisions per km of road*
  - *Collisions per vehicle-km travelled*
  - *Pedal cyclist collisions per km*
  - *Pedal cyclist collisions per cyclist-km travelled*
  - *Pedestrian collisions per km*
- Match this to data on road infrastructure, local environment, and usage provided by Transport for London

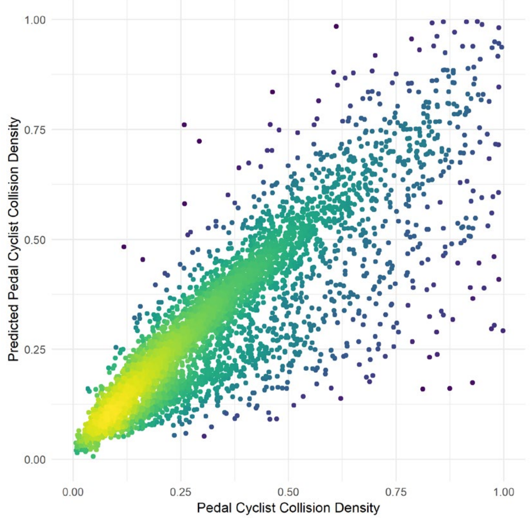
agilysis

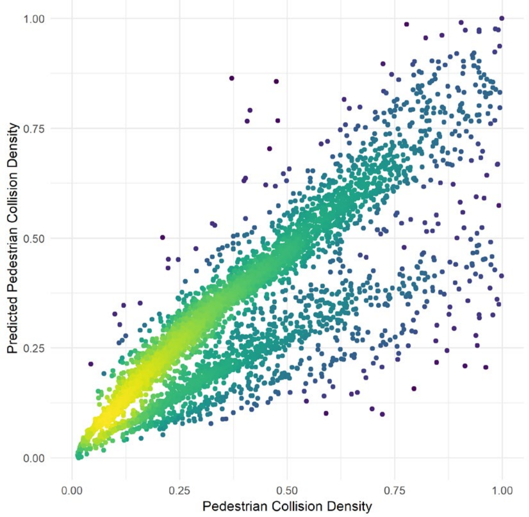
# TFL ROAD DANGER REDUCTION

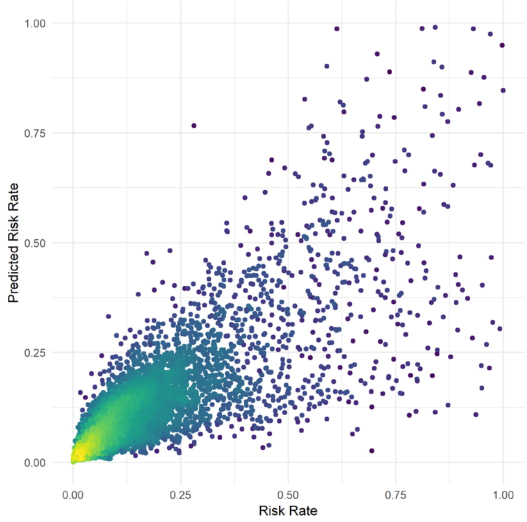
- Train a Neural Network to predict observed measures of risk using this data on road infrastructure, local environment, and usage
- The result is a piecewise-linear function on the space of infrastructure, environment and usage data
- Values for each road segment correlate with observed risk
- The outputs of this function are measures of *Road Danger*

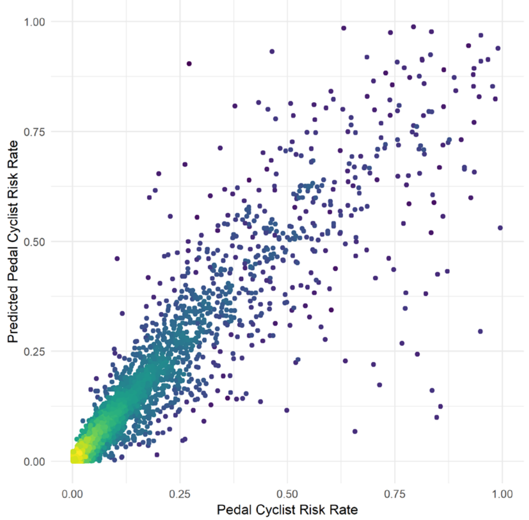
# TFL ROAD DANGER REDUCTION



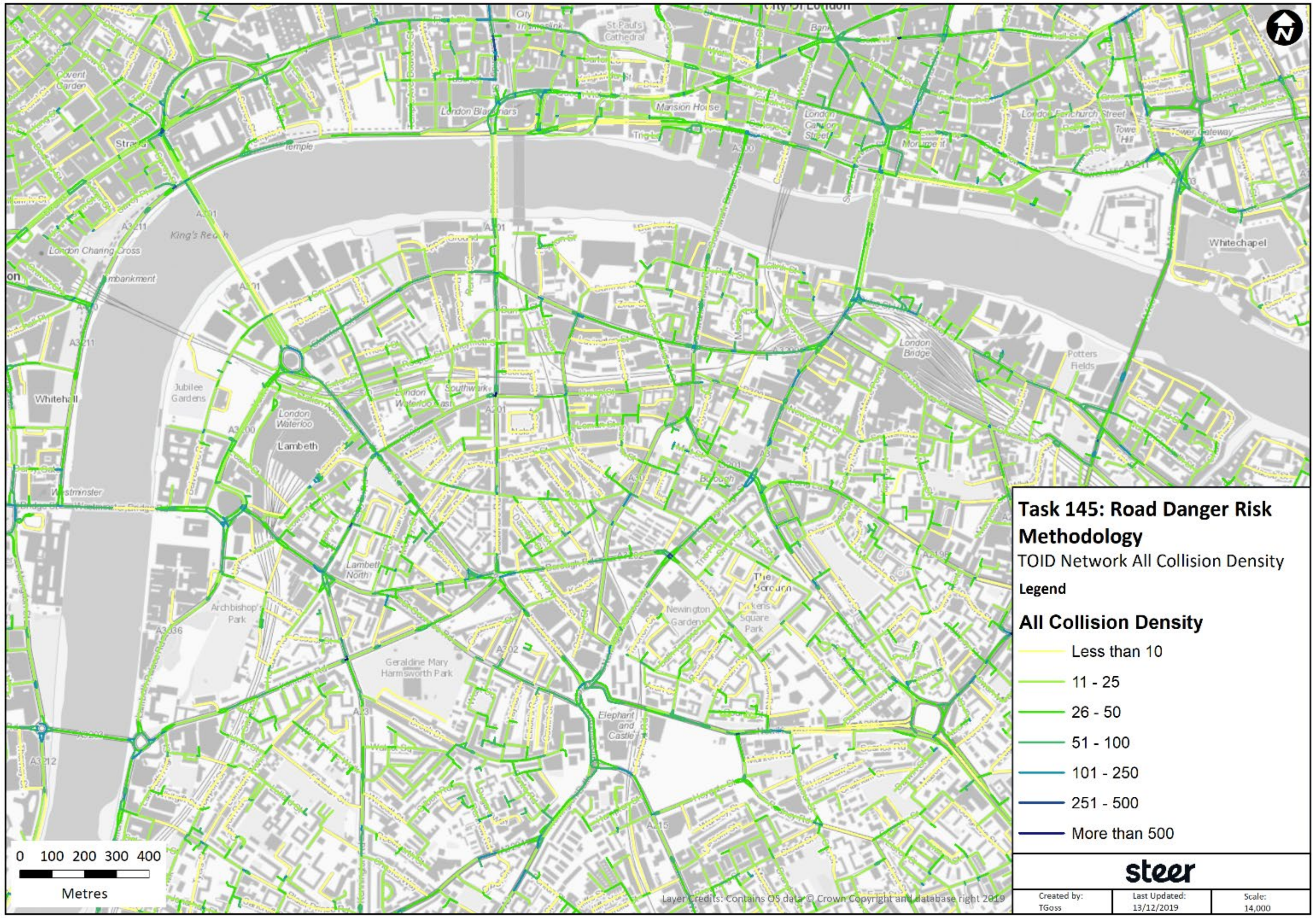




agilysis

# TFL ROAD DANGER REDUCTION

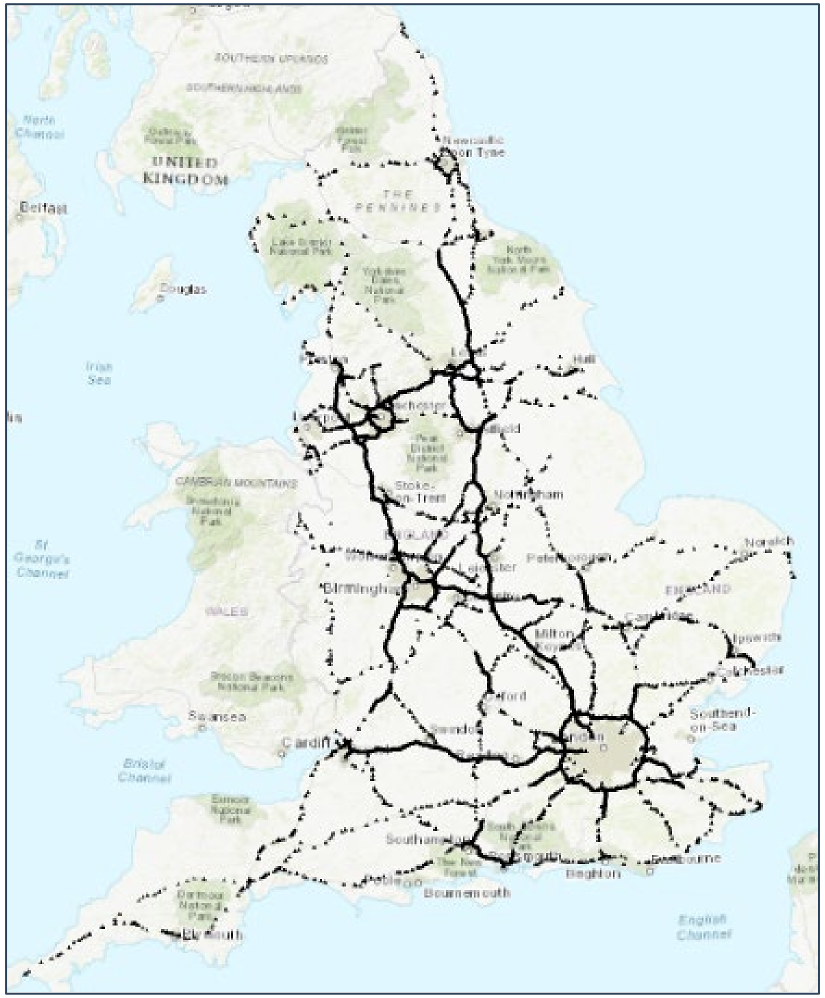



agilysis.co.uk

# WHAT CAN WE USE THEM FOR?

- Regression models
- Classification models

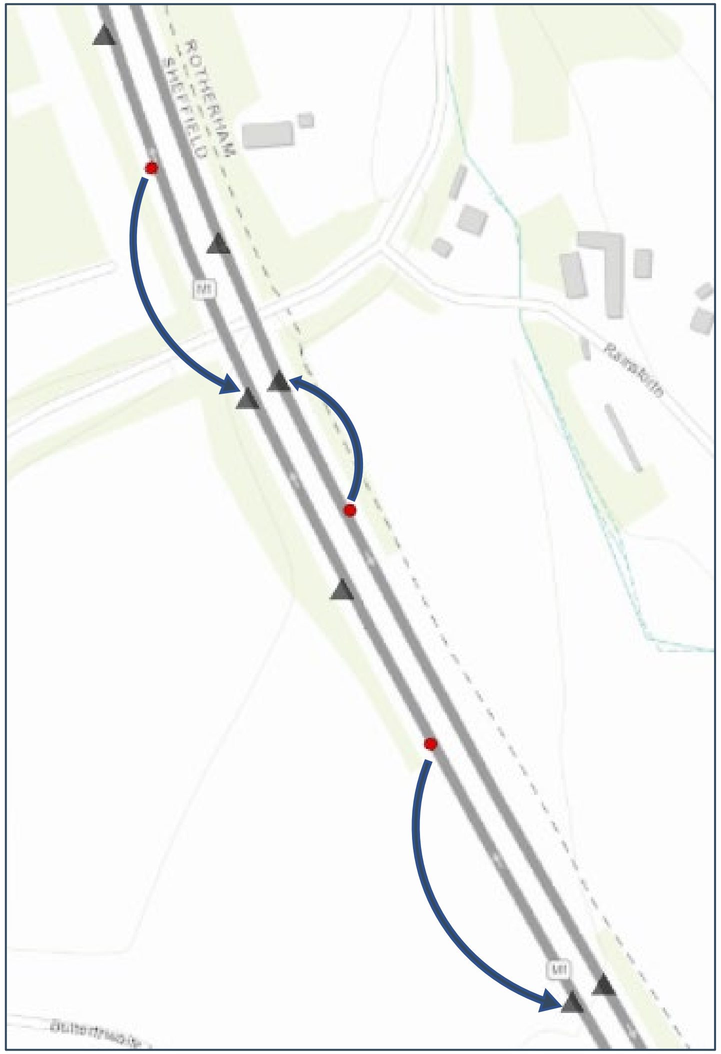
# MIDAS COLLISION DETECTION



- The **MIDAS** network is a collection of sensors across strategic roads in England, collecting minute-by-minute data on speed, flow and headway for each lane
- This project matched these sensors to STATS19 collisions to investigate the extent to which collision effects are visible within MIDAS data

# MIDAS COLLISION DETECTION

- The following parameters were used for selecting collisions
  - **Multi-vehicle** collisions only
  - Occurring **between 6 AM and 10 PM**
  - Unambiguously located on a specific carriageway, not on slip road
- The M1 was chosen as the study site and 2016 as the year
  - Choice was driven by data quality
  - Only 215 (54%) of in-scope 2016 M1 collisions had complete MIDAS data available
- All collisions in scope were matched to the nearest *preceding* MIDAS sensor, to capture changes in vehicle behaviour leading up to the scene of the collision both before and after it occurred



# MIDAS COLLISION DETECTION

agilysis

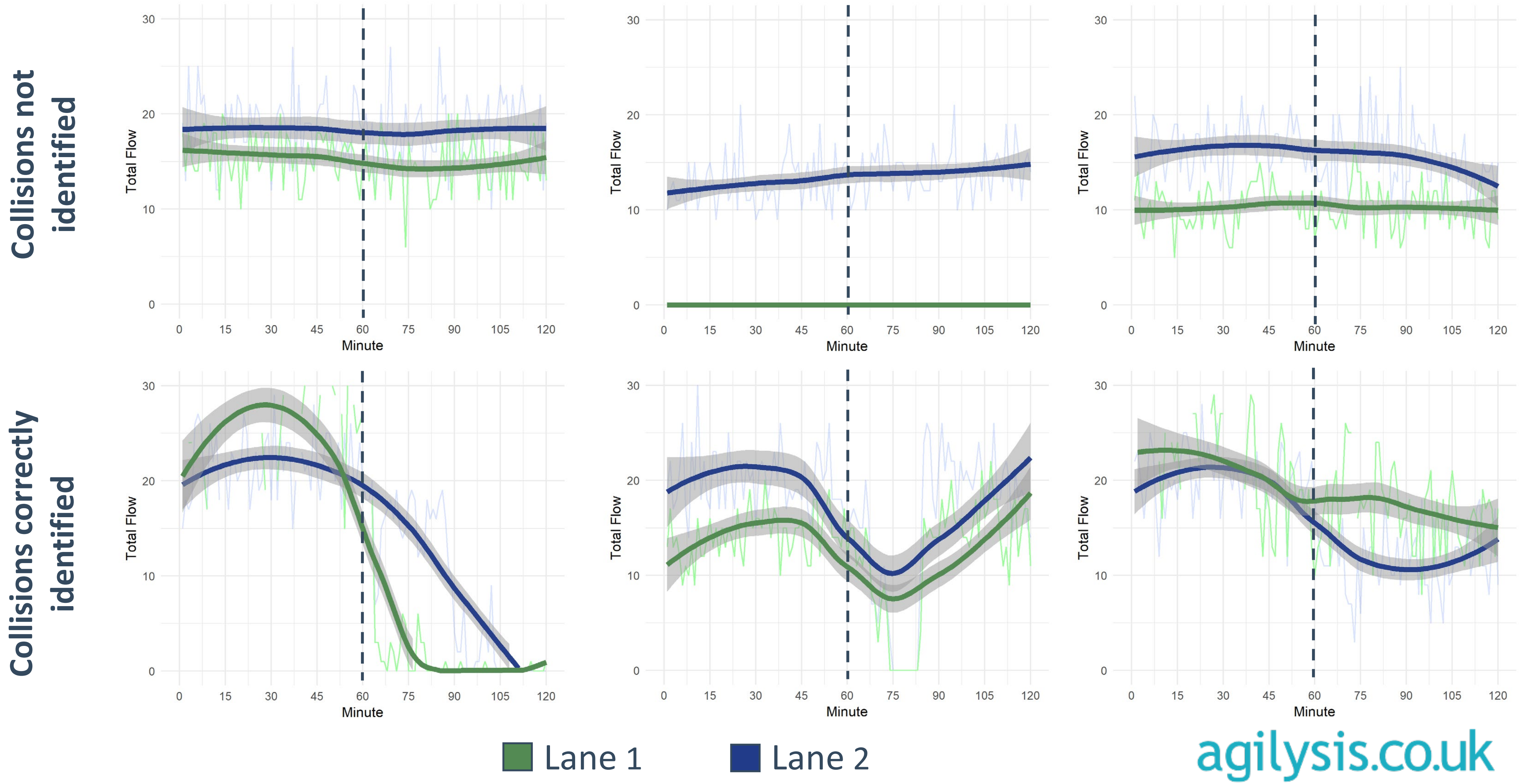
- Multiple two hour periods of MIDAS data were extracted using the RAC Foundation's R package *oneminutetrafficdata*
  - Some were the hour immediately preceding and immediately following a reported collision
  - Others were random comparator periods, between 6 AM and 10 PM on days when no collisions were reported at that location
- Data was only taken for lanes 1 and 2, to account for variation in the number of lanes at each sensor point
- Sensor points are about 300 meters apart on average
- Some datasets were incomplete

# CONCLUSIONS

- The neural network could **accurately recognise** MIDAS data that was not associated with **a collision**, especially using flow data (100% success)
- The neural network was **more likely miss** a collision than it was to identify non-collisions as collisions falsely
- The best factor to use in training a neural network to recognise collisions is **flow** (with a 93% success rate), closely followed by **speed** (which was 89% successful)
- However, training using both seemed **reduce** the model's accuracy
- The neural network **struggled** to recognise patterns in **headway data** that indicated collisions

agilysis

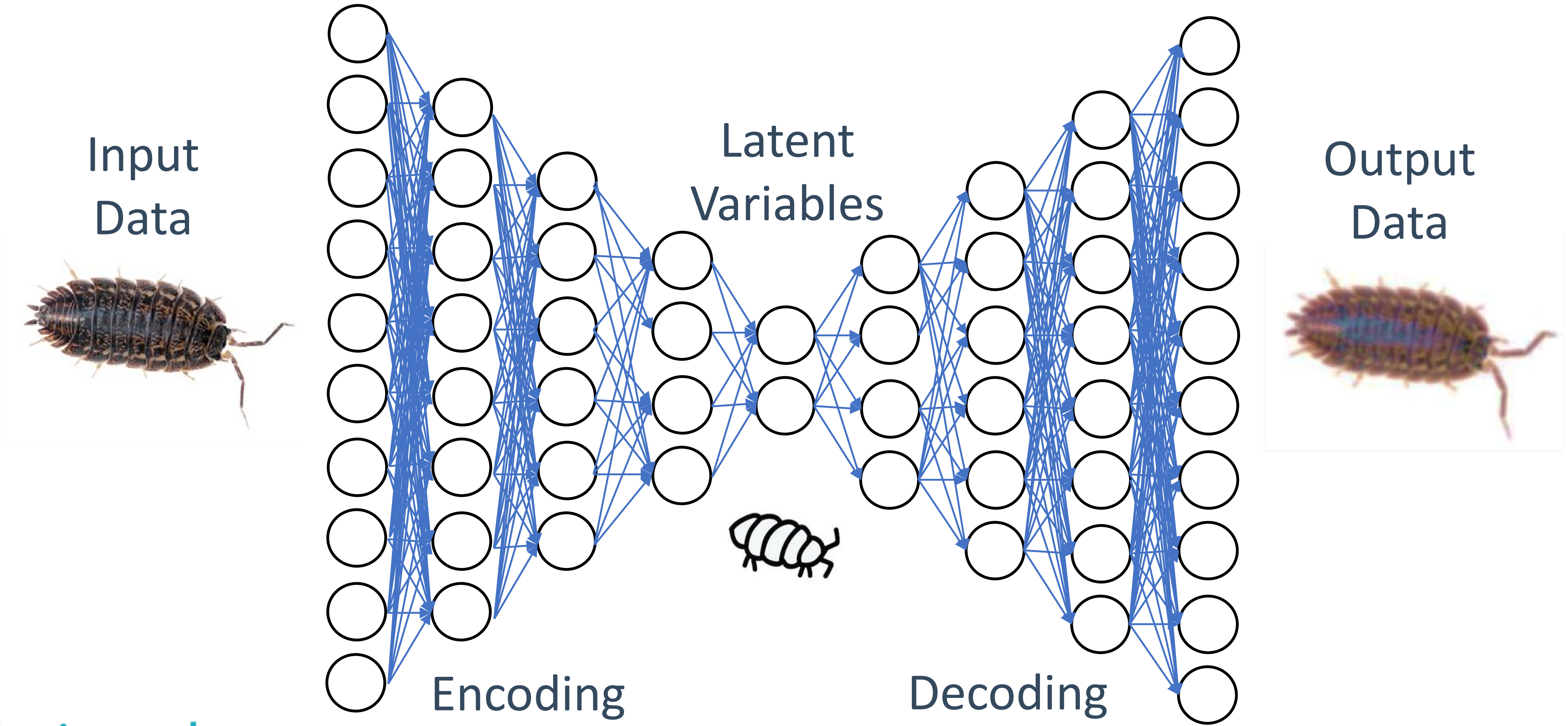
# SAMPLE FLOW PROFILES



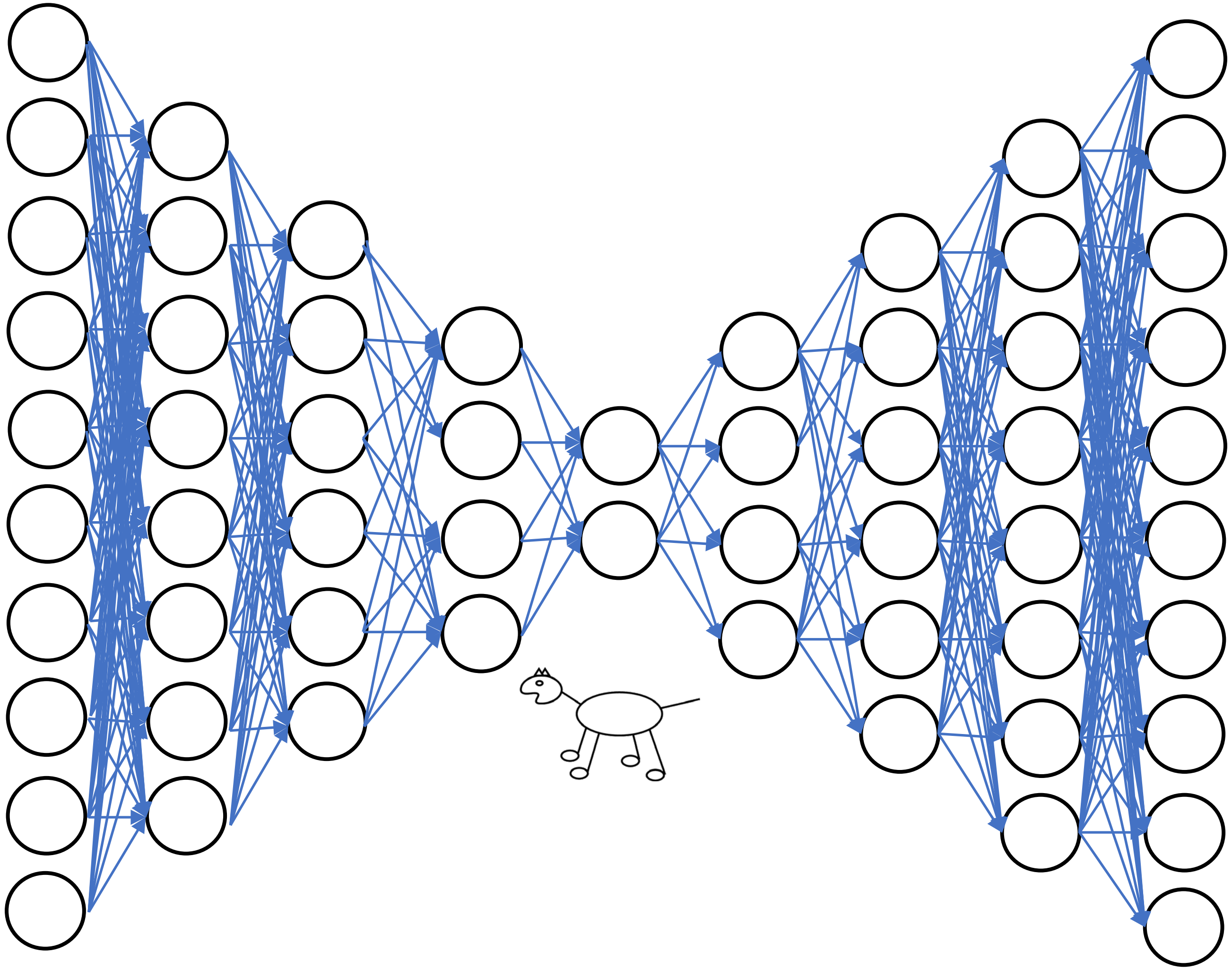
# WHAT CAN WE USE THEM FOR?

- Regression models
- Classification models
- Anomaly/rare event detection

# ANOMALY DETECTION



# ANOMALY DETECTION

≠

# ANOMALY DETECTION

- Exploring the potential of using this to detect the conditions that are likely to precede a collisions in HE's minute-by-minute MIDAS sensor data
- More generally can be used to remove outliers from datasets
- Or as the basis for "unsupervised learning"…

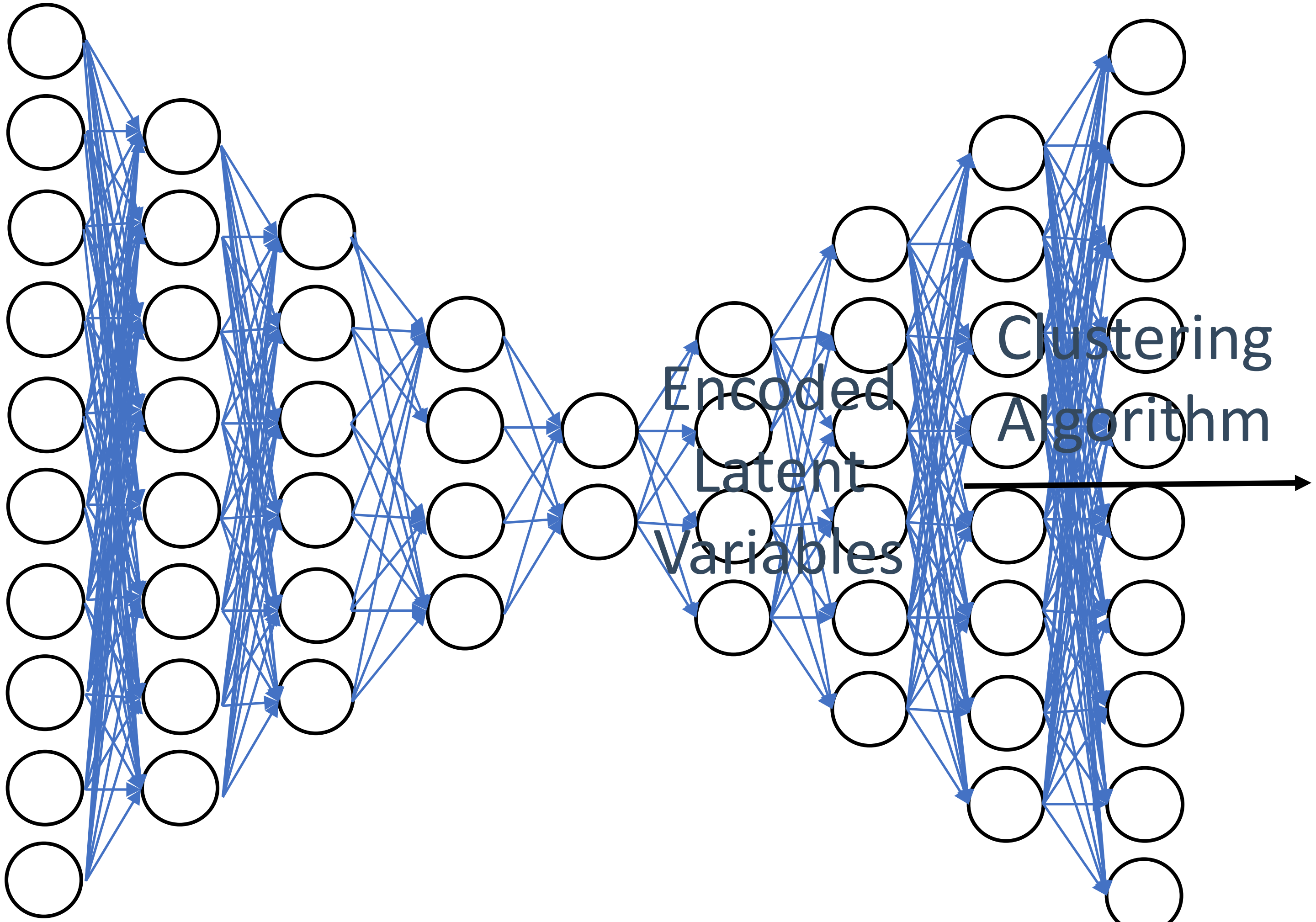
# WHAT CAN WE USE THEM FOR?

- Regression models
- Classification models
- Anomaly/rare event detection
- Unsupervised learning/classification

# UNSUPERVISED LEARNING

agilysis

Input Data

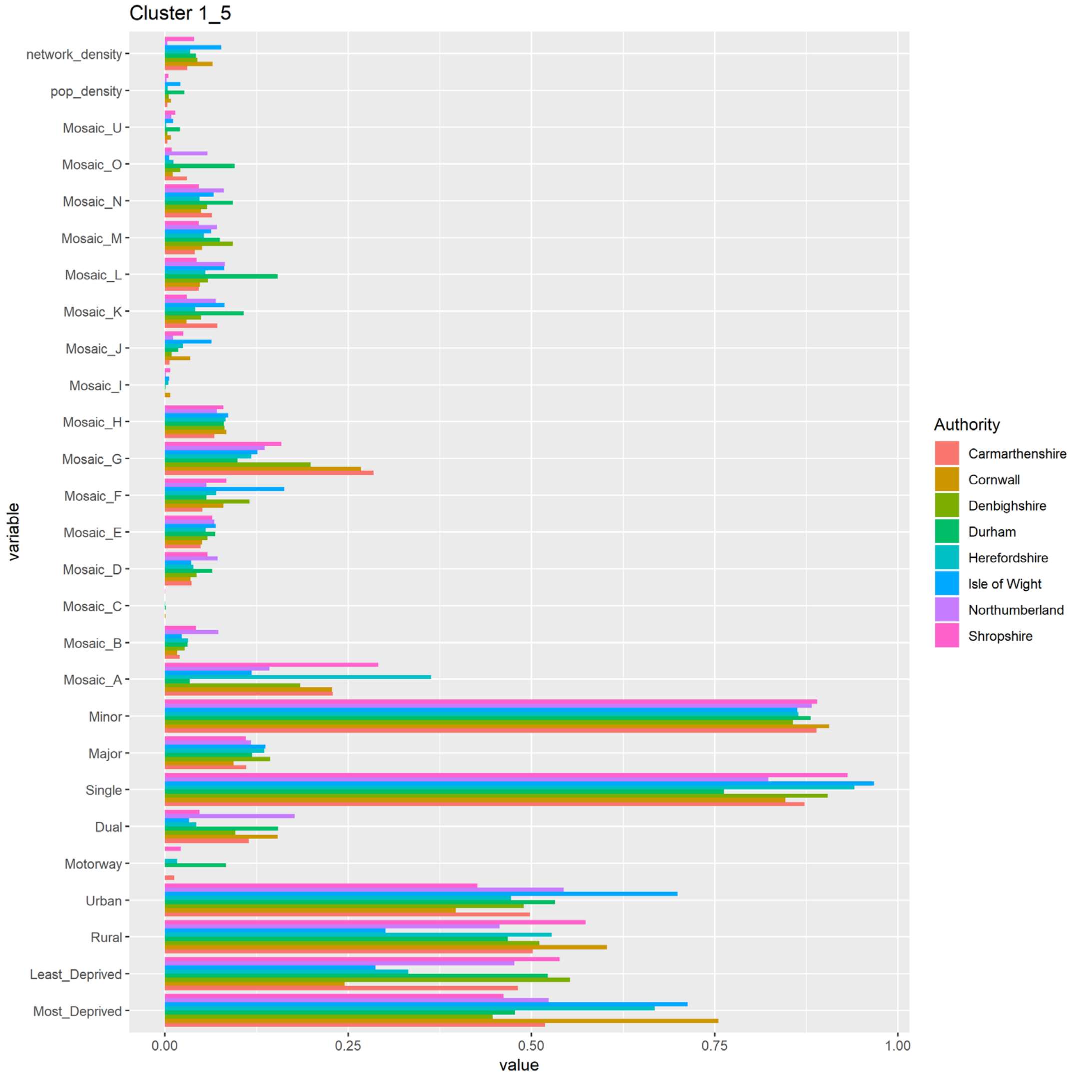


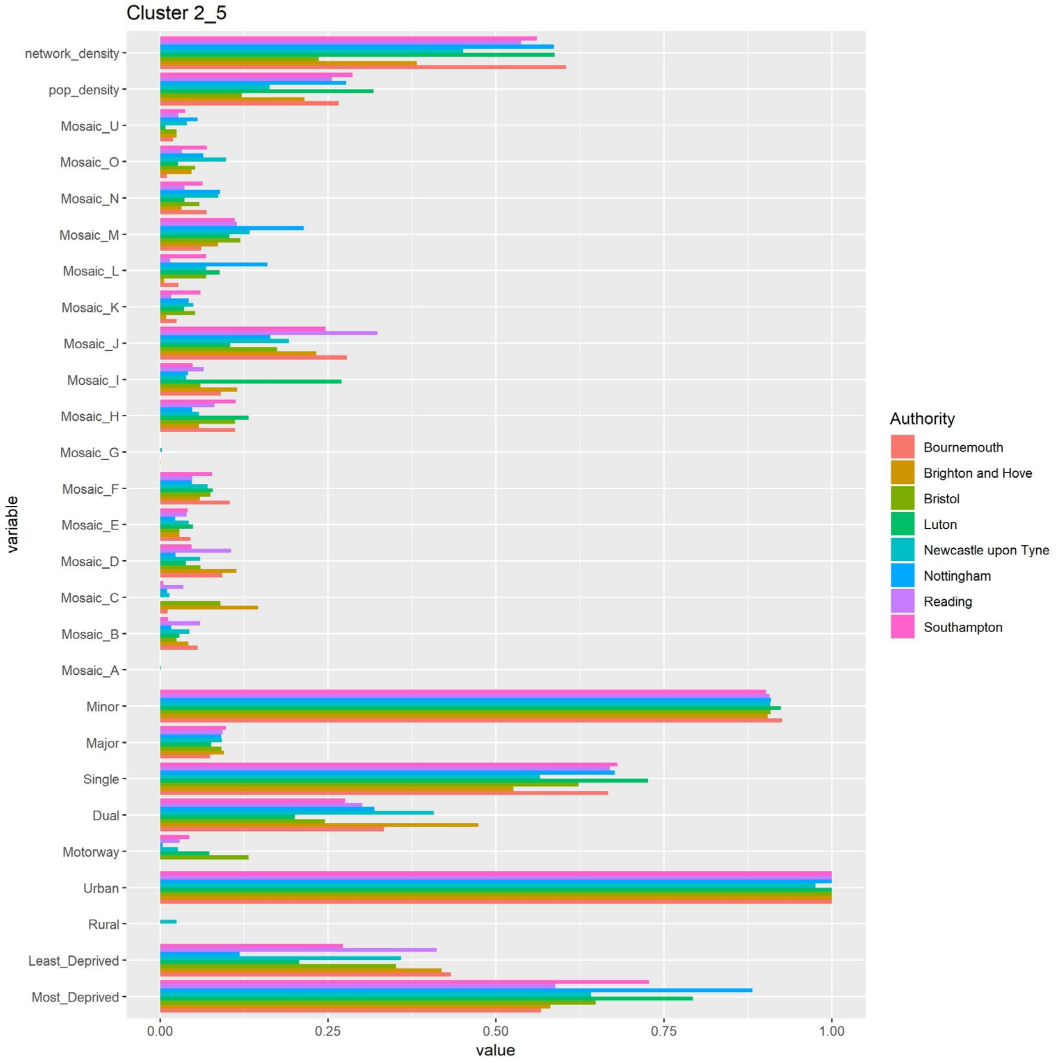

Intelligently
Clustered
Data

# HANCS AND FINDING COMPARATORS

- We work alongside local authorities to analyse their collision data and determine what their biggest road safety issues are
- Only makes sense when compared against *similar* authorities
- Not enough to settle for *similar* meaning nearby
- Similarity should be data driven, and based on factors such as network density, rurality, population density, deprivation levels and socio-demographics
- The HANCS model uses unsupervised learning to cluster authorities into groups based on the level of similarity

# HANCS AND FINDING COMPARATORS



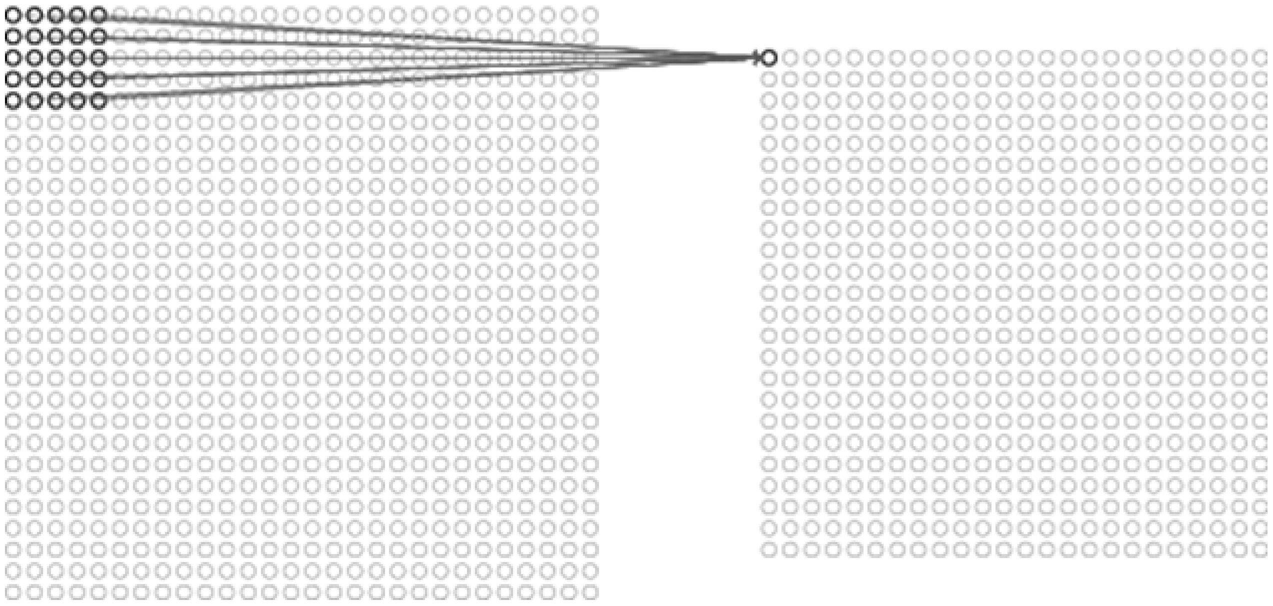



# WHAT CAN WE USE THEM FOR?

- Regression models
- Classification models
- Anomaly/rare event detection
- Unsupervised learning/classification
- Image recognition and object detection

# IMAGE RECOGNITION

## Convolutional Layers



## Pooling Layers

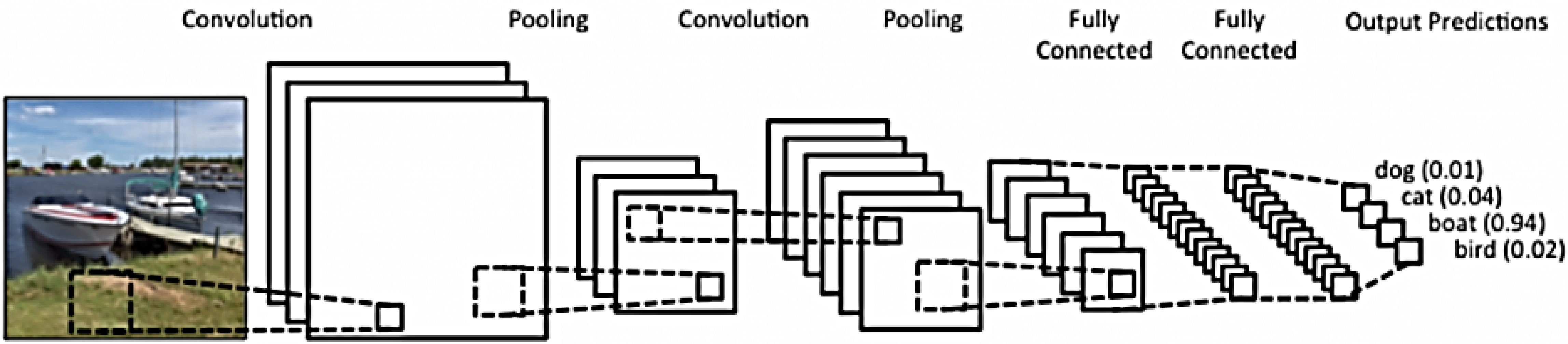
| | | | | | |
|---|---|---|---|---|---|
| 1 | 4 | 7 | 3 | 7 | 9 |
| 2 | 2 | 4 | 7 | 6 | 1 |
| 8 | 6 | 5 | 3 | 4 | 9 |
| 6 | 1 | 7 | 5 | 9 | 3 |
| 6 | 4 | 5 | 2 | 3 | 2 |
| 9 | 6 | 4 | 5 | 8 | 6 |

→

| | | |
|---|---|---|
| 4 | 7 | 9 |
| 8 | 7 | 9 |
| 9 | 5 | 8 |

# IMAGE RECOGNITION

agilysis



# IMAGE RECOGNITION

# RAPIER

- Road assessments can be hugely beneficial in determining how best to create a safe system
- A lot of these require manual coding of road features, which can be costly and time consuming
- This limits impact
- Long term goal to use machine learning techniques to programmatically code road infrastructure on the network using a combination of satellite imagery and video recorded by vehicles
- Used in combination with models that determine risk levels based on the coded infrastructure

# RAPIER

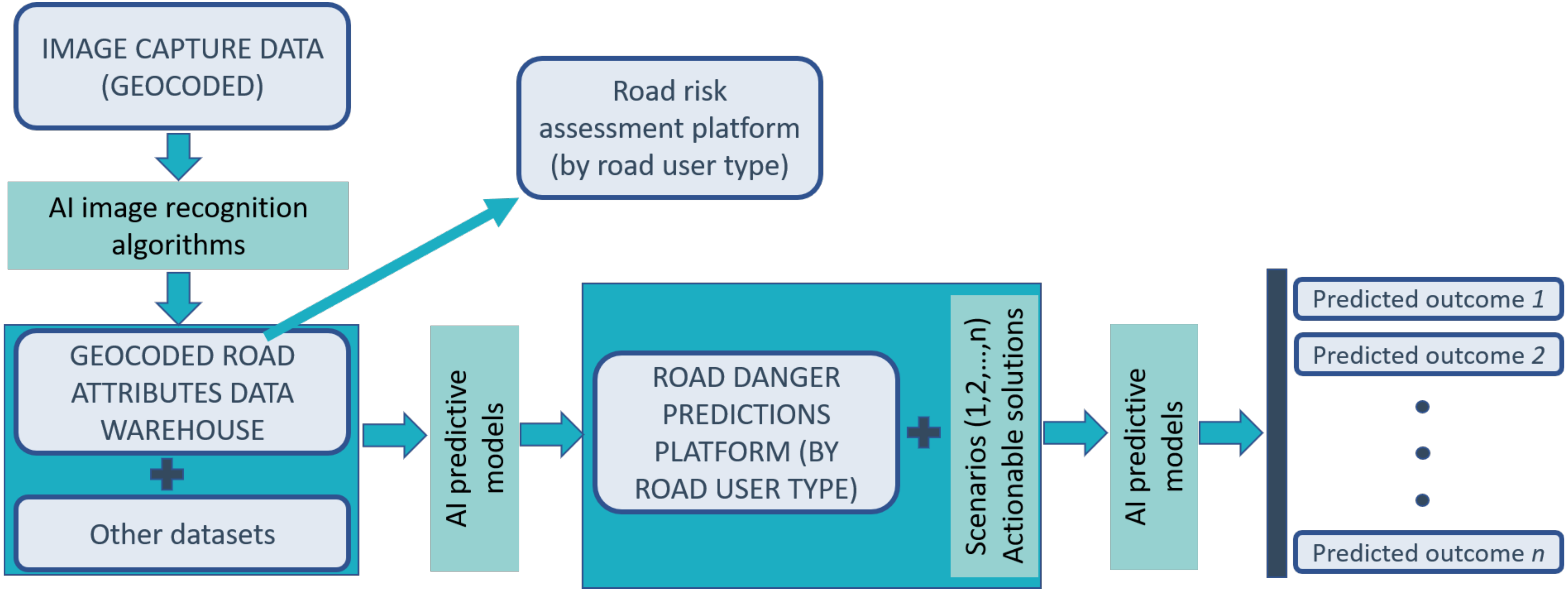


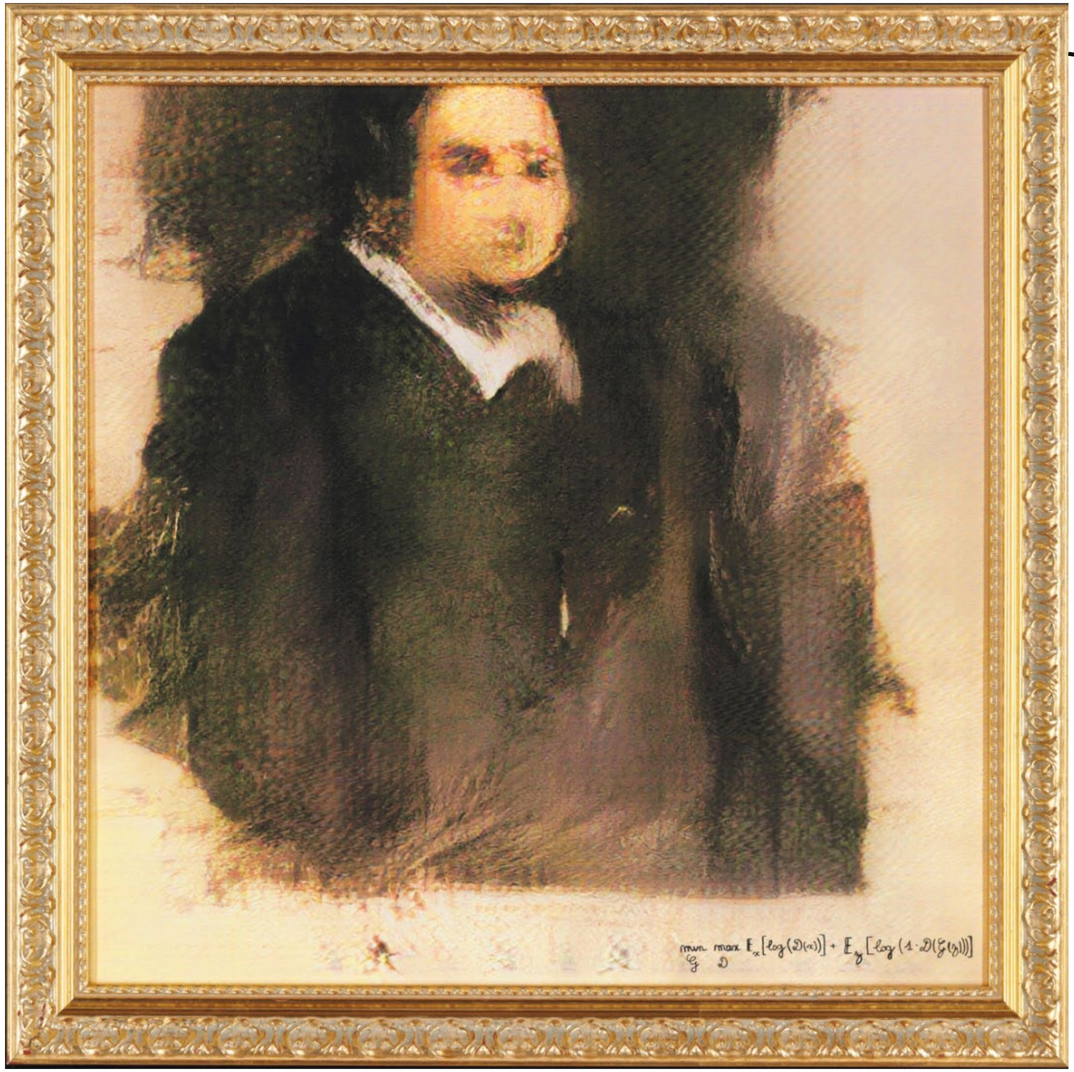
IMAGE CAPTURE DATA (GEOCODED)
AI image recognition algorithms
Road risk assessment platform (by road user type)
GEOCODED ROAD ATTRIBUTES DATA WAREHOUSE
Other datasets
AI predictive models
ROAD DANGER PREDICTIONS PLATFORM (BY ROAD USER TYPE)
Scenarios (1,2,…,n)
Actionable solutions
AI predictive models
Predicted outcome 1
Predicted outcome 2
Predicted outcome n

# WHAT CAN WE USE THEM FOR?

- Regression models
- Classification models
- Anomaly/rare event detection
- Unsupervised learning/classification
- Image recognition and object detection
- Art

# AI ART

agilysis



$432,500
Christie's, 2018

# DEEP DREAMING

More dog

agilysis


# DR CRAIG SMITH
# BRUCE WALTON

+44 1295 731810

craig.smith@agilysis.co.uk

Bruce.Walton@agilysis.co.uk