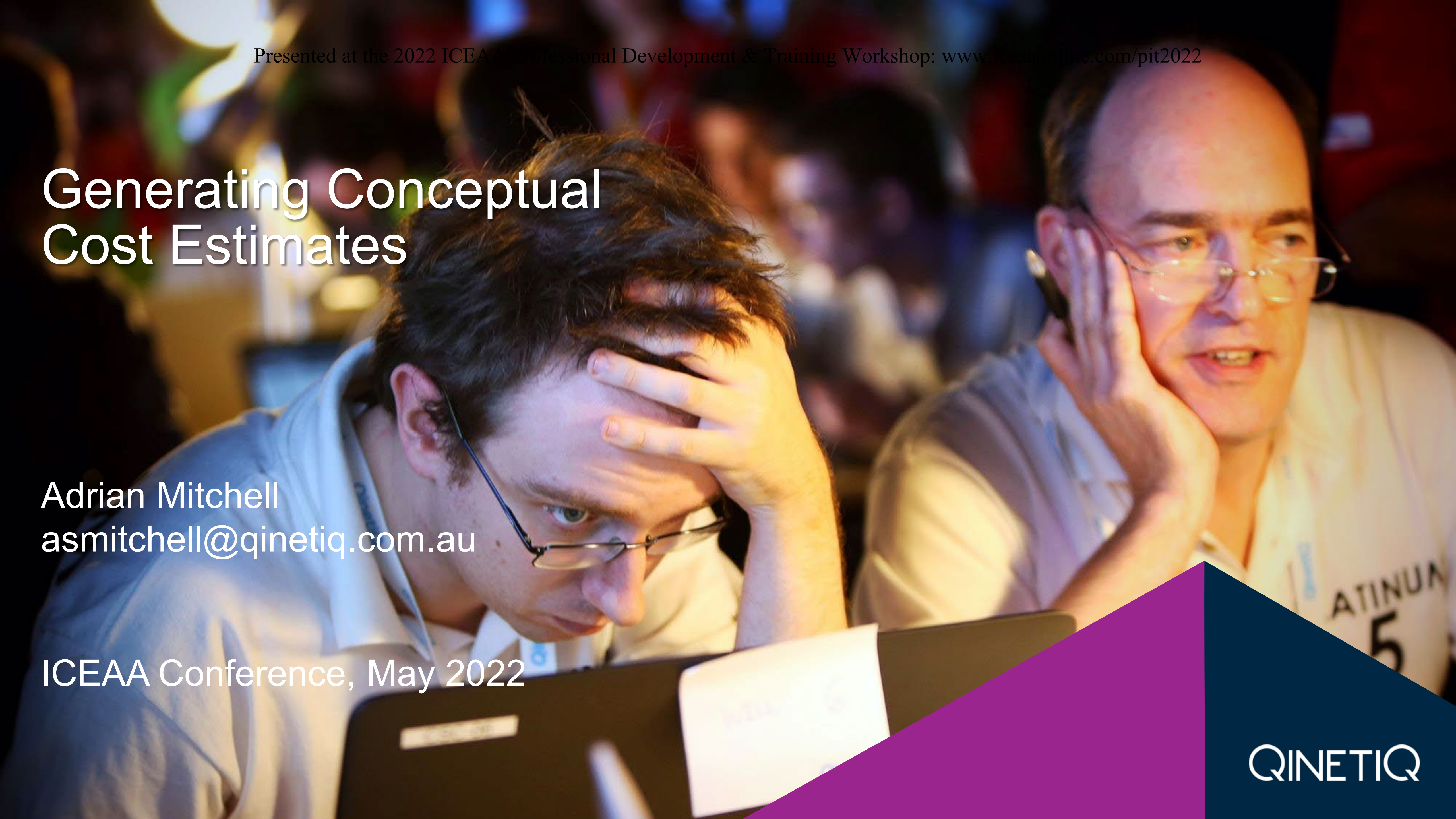


Generating Conceptual Cost Estimates

Adrian Mitchell
asmitchell@qinetiq.com.au

ICEAA Conference, May 2022

QINETIQ

Background

Opportunity

- Opportunity
  - Generate tens of new cost estimates for ideas looking to become projects.
    - Limited idea of requirements and project scope
    - When time and resources are limited
    - The need to withstand external scrutiny
    - Requirement for total ownership cost estimates
  - Facilities to support mission systems are often not the focus of cost practitioners.
  - The cost of facilities is sometimes overlooked when setting an early project budget.

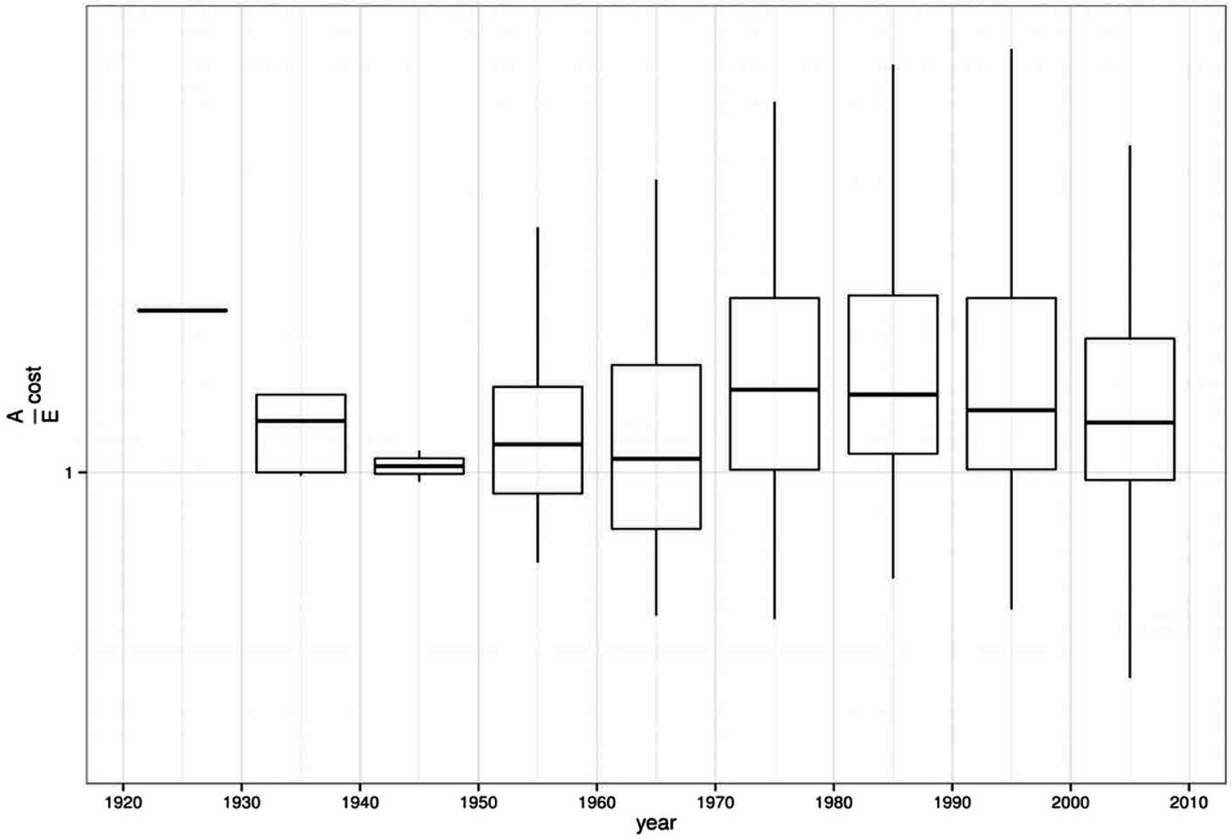
- Challenge
  - Early estimates are incredibly inaccurate
  - Traditional bottom-up costing and linear regression methods, by themselves, may struggle
  - Complex models hurt more than they help

- Solution
  - A software driven process
  - Thousands of rows of trained and tested data
  - Multiple prediction methods

Presented at the 2022 ICEAA Professional Development & Training Workshop: www.iceaaonline.com/pit2022

Challenge

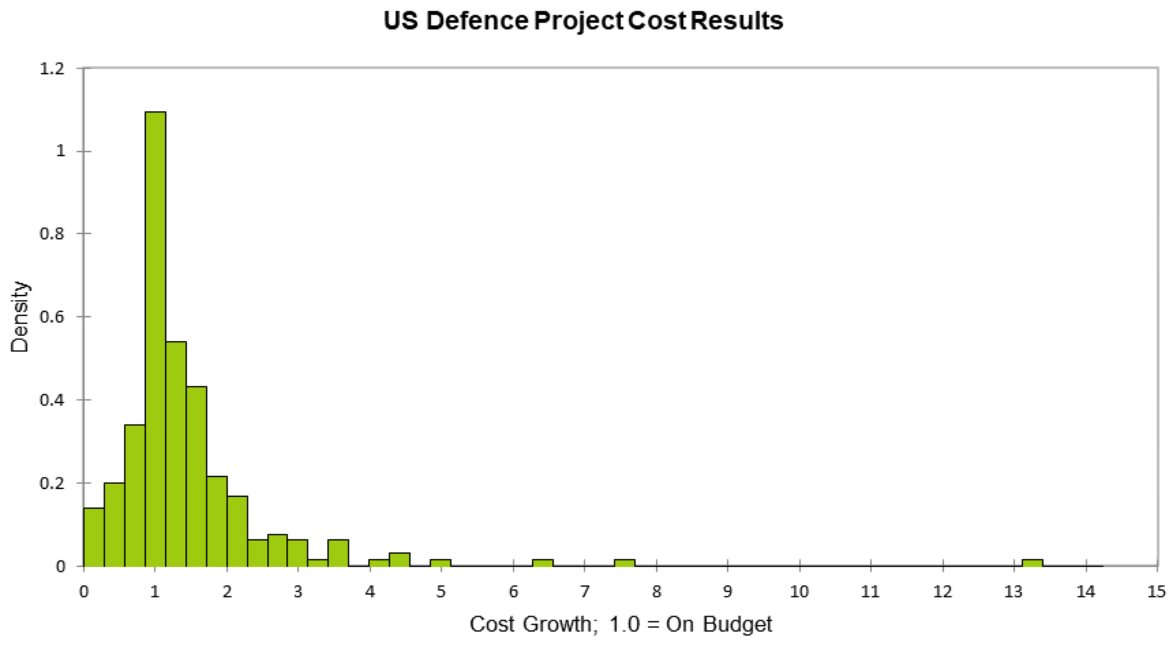
- On average civil infrastructure projects are subject to an overrun of 39 per cent. The international trend since the 1950s has remained similar (n = 1603, p<0.0001). Bent Flyvbjerg, The Fallacy of Beneficial Ignorance: A Test of Hirschman's Hiding Hand, World Development Vol. 84, pp. 176–189, 2016.
- For 'mega' civil infrastructure projects: "Costs are underestimated in almost 9 out of 10 projects. For a randomly selected project, the likelihood of actual costs being larger than estimated costs is 86%." (n = 258, p<0.001). Bent Flyvbjerg, Mette Skamris Holm, and Søren Buhl, Underestimating Costs in Public Works Projects: Error or Lie?, Journal of the American Planning Association, vol. 68, no. 3, Summer 2002, pp. 279-295.



COMMERCIAL IN CONFIDENCE

Challenge

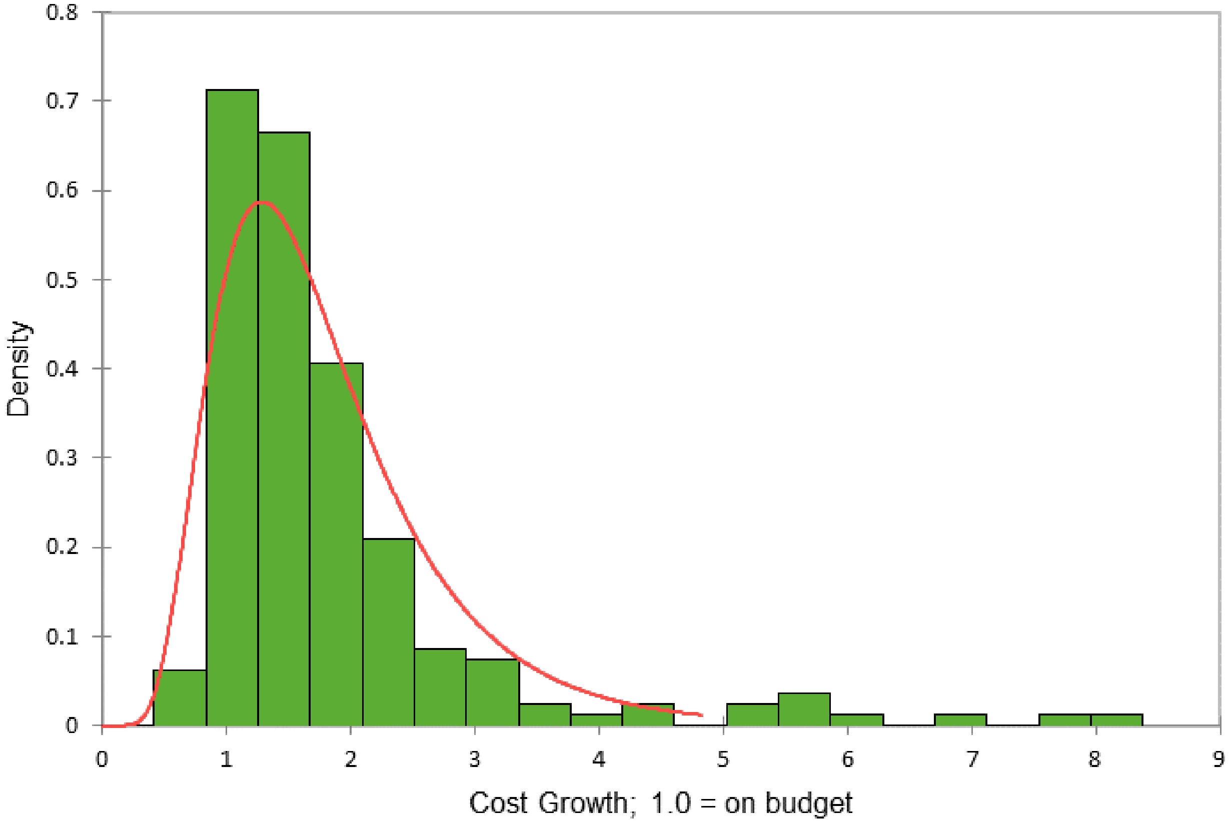
- Public Selected Acquisition Report data for 225 planning and development projects
  - Normalised for production quantities
- The probability of completion on or below the original project cost forecast is 35 per cent.
- The probability of a project coming within 1 per cent of actuals is 6 per cent.



QINETIQ

Challenge

- If cost overruns were only due to estimating errors, then according to Central Limit Theorem the differences between actual and estimate results should tend towards a normal distribution.
- ….instead the results tend towards a distribution that is heavily skewed, suggesting that behavioural biases produce overly optimistic project cost estimates.



- International civil and defence (non-QinetiQ) projects from QinetiQ's database
- n = 195, p-value <0.0001.

EstatiQ

Solution

- EstatiQ (Estate + QinetiQ) generates conceptual cost and duration estimates of facilities proposals.
- The model is a supervised ensemble of artificial intelligence methods called a 'super learner', trained and tested on 125,000 rows of Department of Defence and industry data.
- Generating an estimate takes a few minutes.
- EstatiQ uses python open-source technologies.
- The super learner is demonstrably more accurate than any individual prediction method.
  - Evidence from academia and industry shows that conceptual cost estimates generated for complex projects using accounting or quantity surveyor methods are extraordinarily inaccurate.
  - Considering the lead times for Defence projects, there is little point spending large resources generating estimates that fail to approximate actual cost - despite the best efforts of the cost practitioner.

Scope of the Model – Product Breakdown Structure

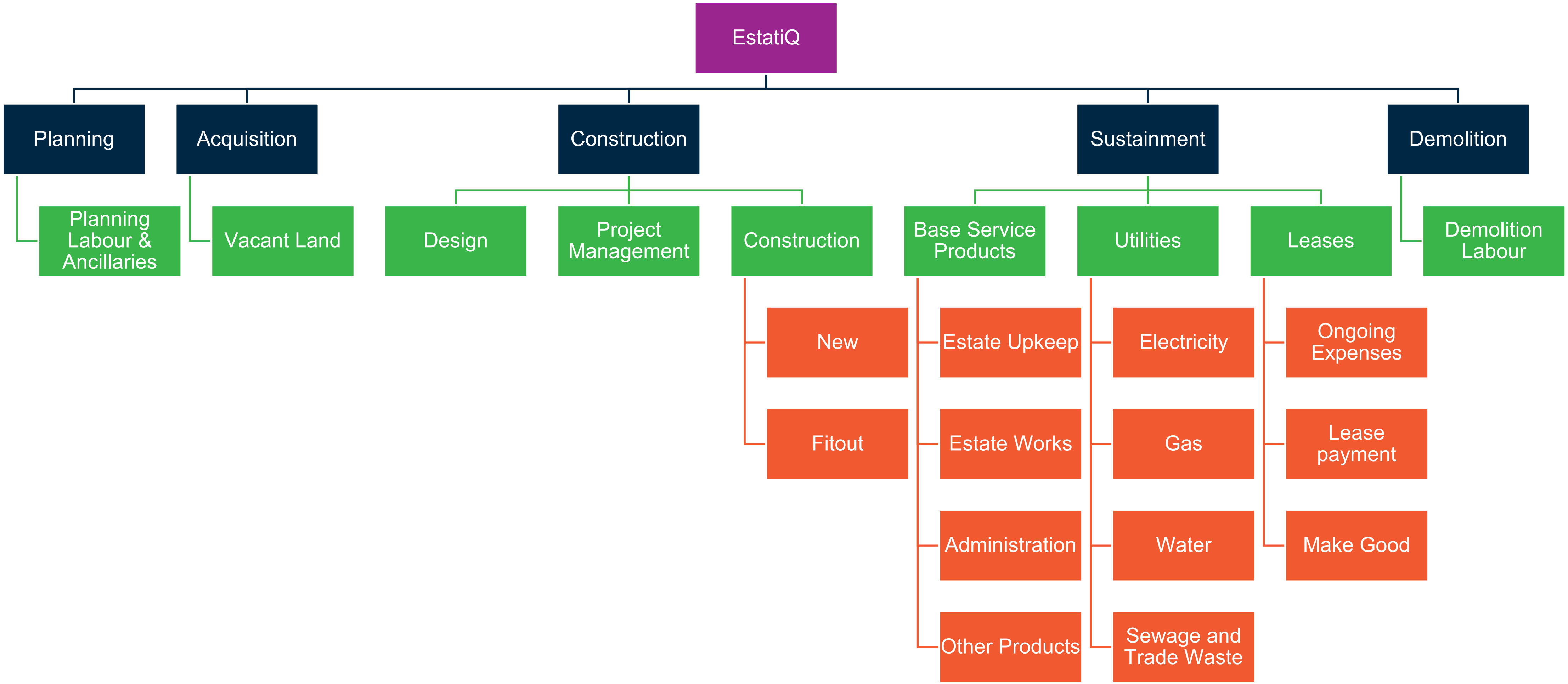
Property

Land

Buildings

Infrastructure

Scope of the Model - Breakdown Structure

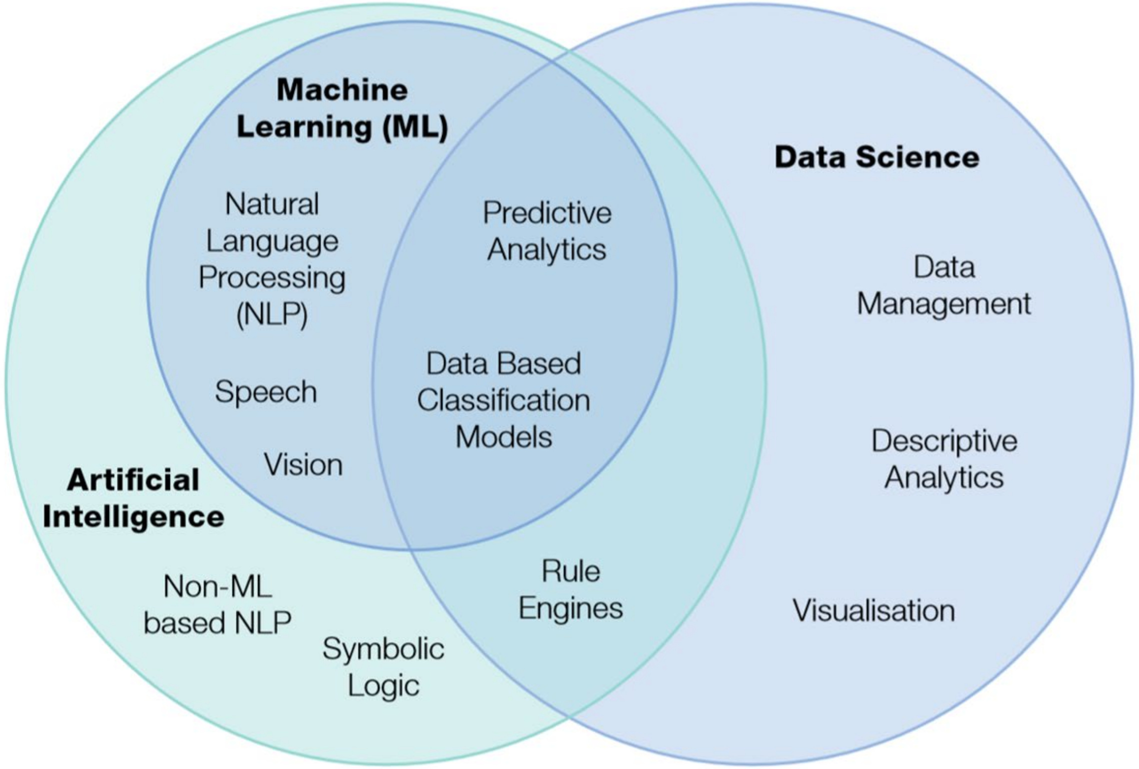


QINETIQ

Methods

What is Machine Learning?

- …a branch of artificial intelligence that draws on areas of computer science, mathematics and statistics
- Machine learning is concerned with improving the ability of a machine to complete a certain task in response to experience.
- Such tasks usually involve:
  - Categorising data
  - Prediction
  - Decision making
  - Optimisation
  - Pattern recognition
- Algorithms are trained on data and over time the algorithms learn from new data to make more accurate predictions.



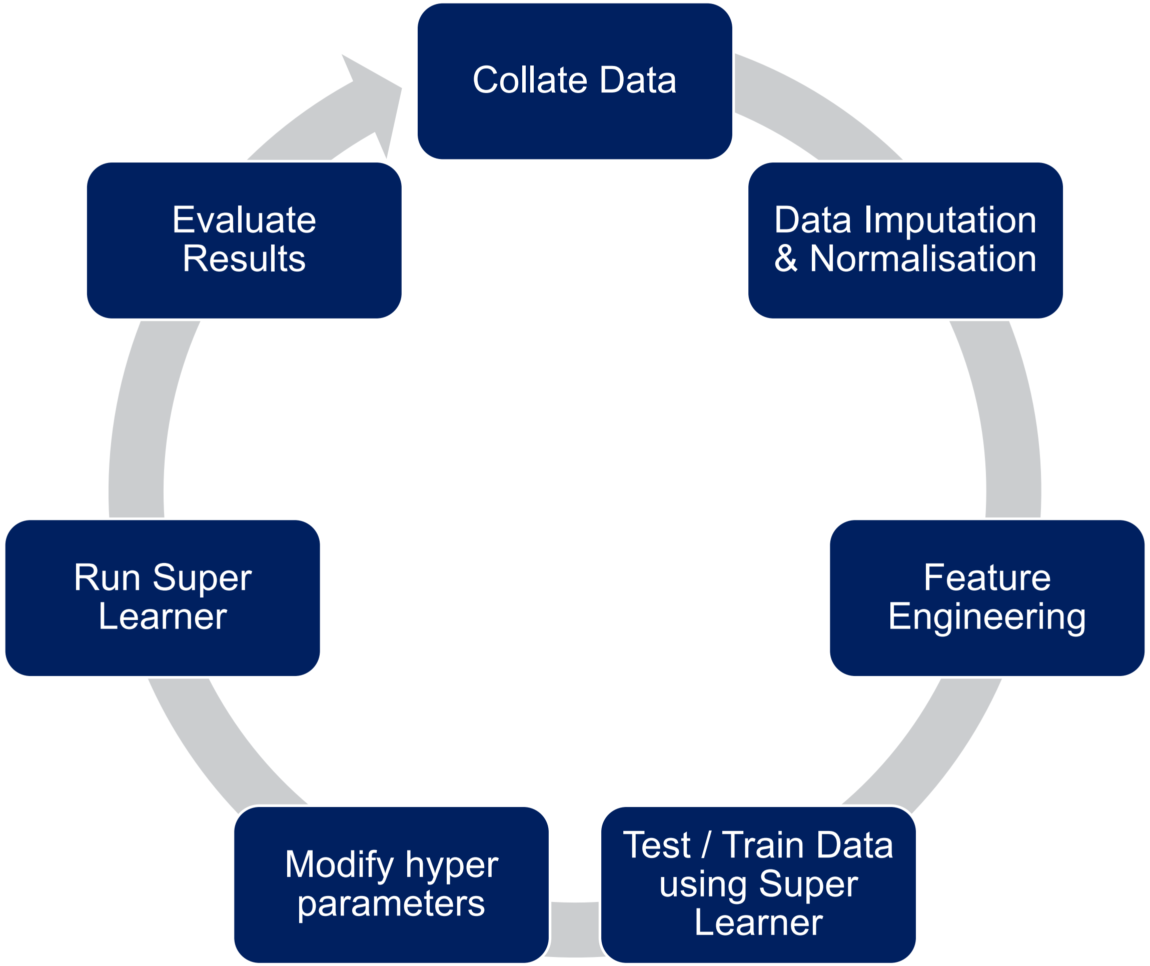
COMMERCIAL IN CONFIDENCE
QINETIQ

Why use Machine Learning?

Can the Method Model?	ML	Build-Up
Complex, massive, messy datasets	Yes	No
…when relationships are not clear	Yes	No
…despite missing values	Yes	No
Qualitative data	Yes	No
Missing categorical values	Yes	No
Non-linear relationships	Yes	No
Forecast residual and error results	Yes	No
Small data sample (rows) with many predictors (columns)	Yes	Yes, but not accurate
Quick turnaround estimates	Yes	No

- Regression: predict a quantitative response variable: "What should the building cost?"
- Classification: predict a qualitative response variable. "What material should we use to construct the building?"
- The potential number of combinations that could be used to calculate the construction cost for a single Defence building is in the nonillions.

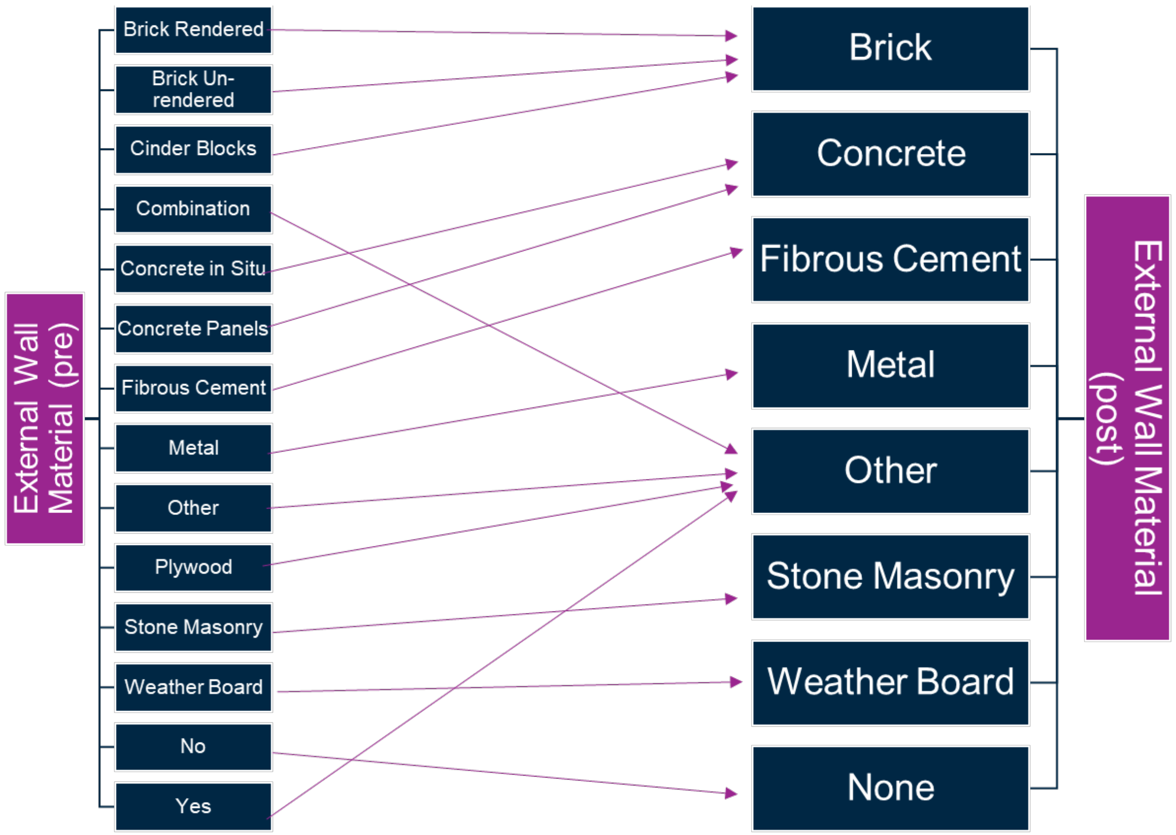
Modelling Process



Feature Engineering

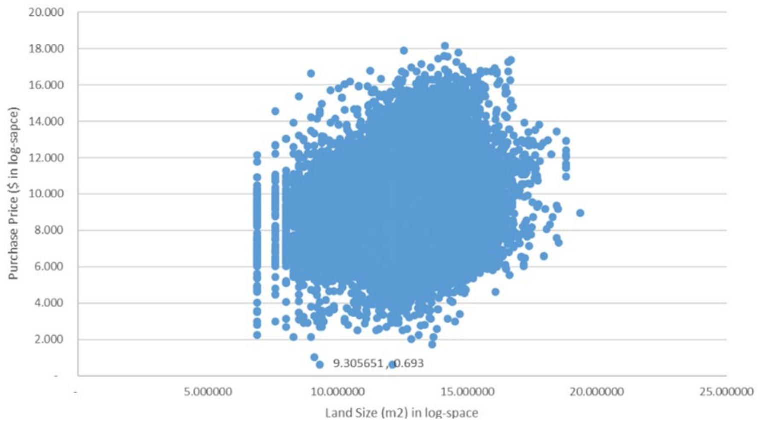
Feature Engineering

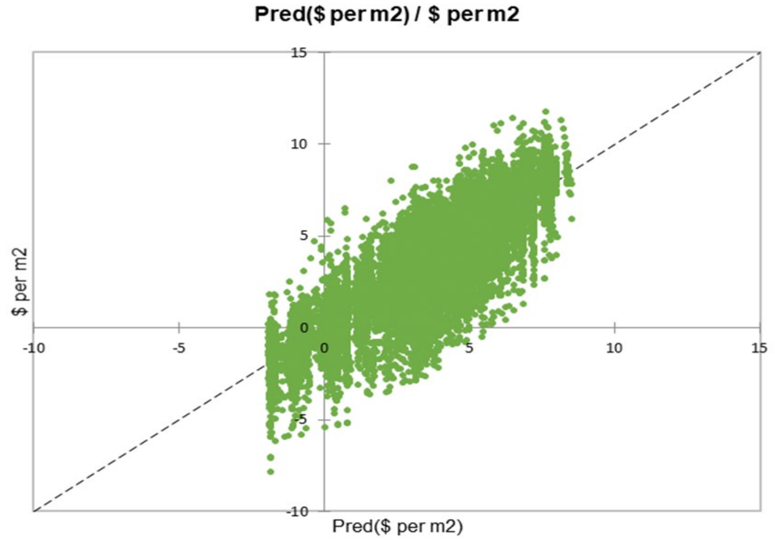
- Enrich data sets to improve model predictability, balanced against model usability by consolidating, converting or creating data.
- Consolidate qualitative features
  - Mitigate data quality issues
  - The meaning behind some categories is unclear
  - Some categories are too sparsely populated to yield meaningful statistical results
  - Align features with the early conceptual nature of EstatiQ's estimates



Feature Engineering

- Convert quantitative data to qualitative data
  - Number of floors were replaced with labels Single-Storey and Multi-Storey, to overcome data quality issues and the 'curse of dimensionality'.
  - Land size was converted to labels to improve model predictability and dollars to $ per m2.
    - Vast > 1,000,000 m2
    - Large > 100,000 m2
    - Medium > 1,000 m2
    - Small > 100 m2
    - Tiny < 100 m2





Feature Engineering

- Create datasets
  - Infrastructure is a broad category that encompasses up to 46 seemingly unrelated sub-categories of different asset types.
  - EstatiQ uses variables that are common to all infrastructure sub-categories to 'pull together' the dataset for running ensemble methods.
  - Profile of the asset, material and then complexity in relation to all other assets of the same profile.
  - For example, a runway has a surface profile, is made of composite rock and is complex, while a surface playing field is made of earth and is basic.

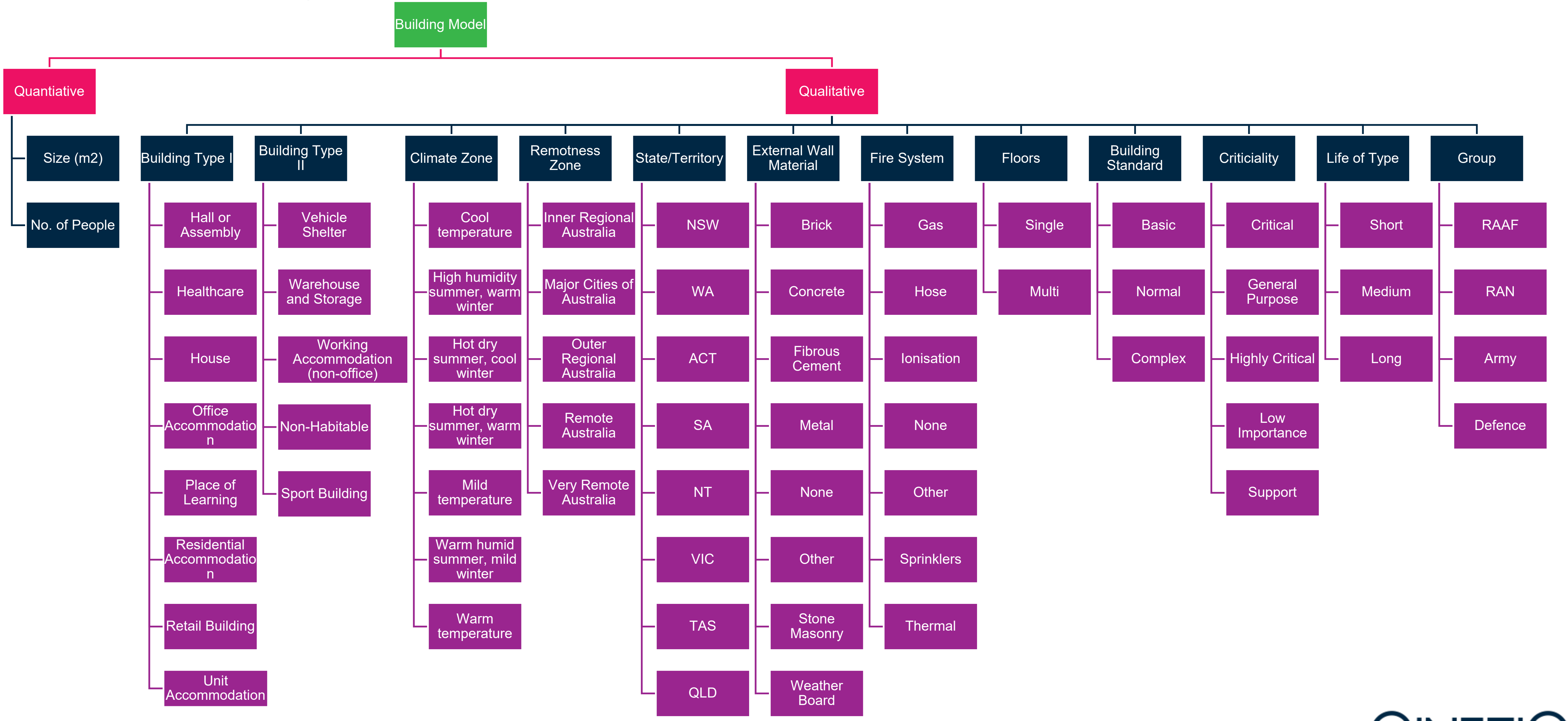
Profile	Primary Material	Complexity
Elevated	Composite Rock	Basic
Subterranean	Earth	Normal
Surface	Metal	Complex
	Plastic	
	Wood	

Feature Engineering

- Create datasets…continued
  - The infrastructure data set lacked sufficient sizing data to support the implementation of a plausible data imputation strategy.
  - A categorical sizing variable is created from residuals generated during the testing and training of the model.
  - The sizing labels (Large to Small) are created and then checked against the few sizing data observations that do exist, to create handy lookup tables for the end-user.
  - Various rounds of testing and training models were required to provide a satisfactory result.
  - This approach aligns infrastructure with the early conceptual nature of the EstatiQ model.
    - At an early stage, will a cost practitioner know the exact length of all the electrical conduits required for a new Army base?
    - The cost practitioner will know that an electrical distribution system is required for an Army base in a certain location.
    - A base wide electrical distribution system is then selected as size “Large” by the user in EstatiQ.

COMMERCIAL IN CONFIDENCE

Example: Building Model Features



COMMERCIAL IN CONFIDENCE
QINETIQ

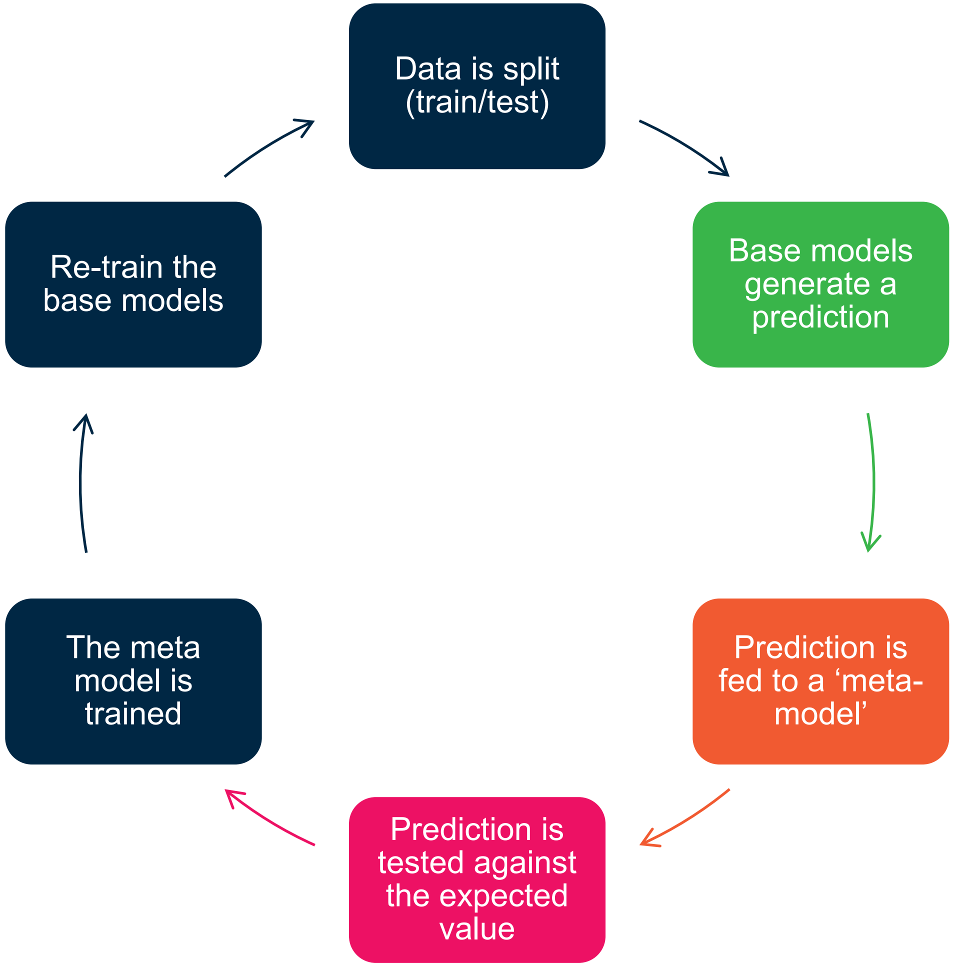
Super Learner

Presented at the 2022 ICEAA Professional Development & Training Workshop: www.iceaaonline.com/pit2022

Super Learner

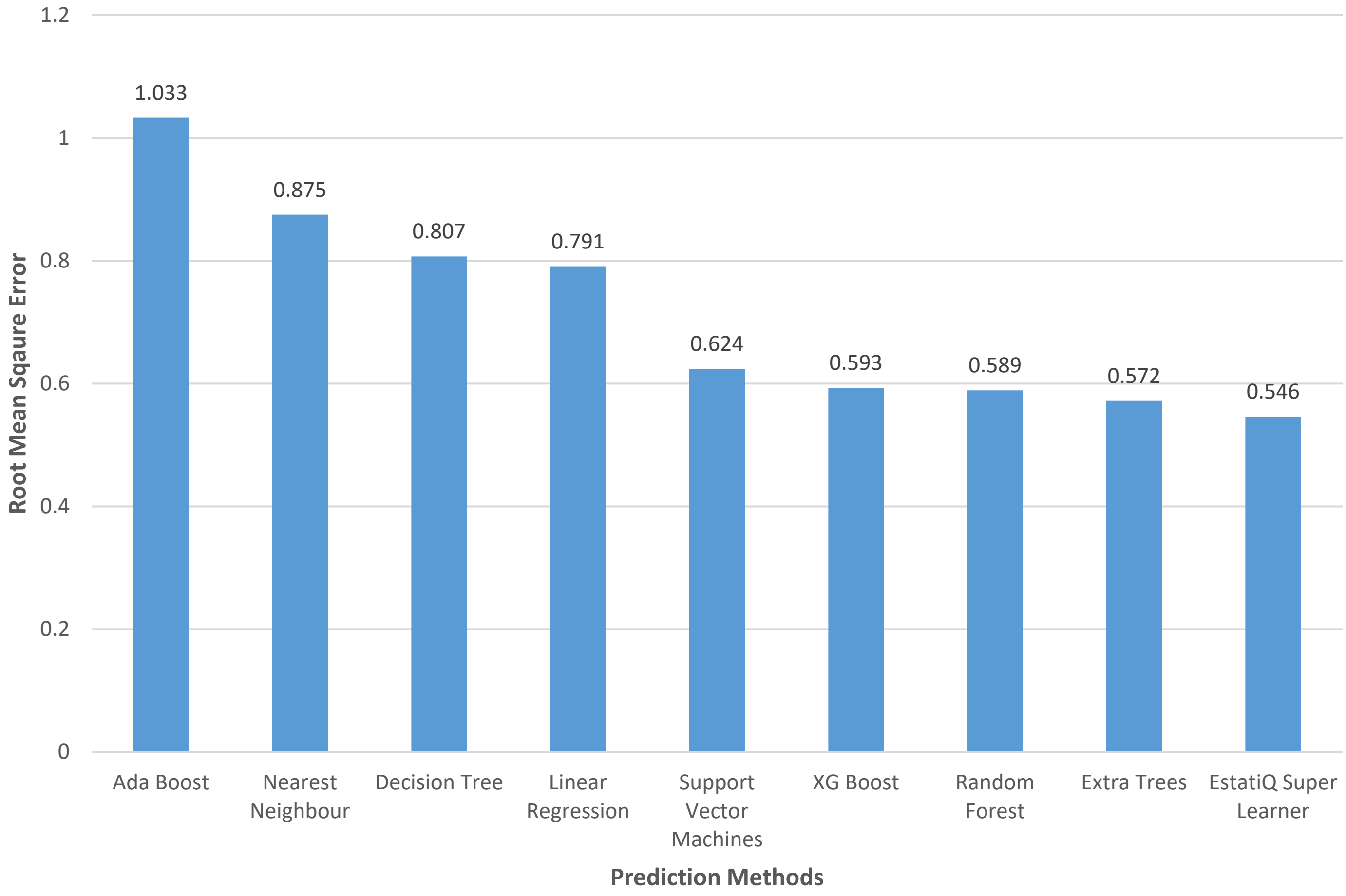
- EstatiQ is an ensemble of prediction methods, created using a process called stacking.
- Individual methods generate a prediction and are tested using common 'meta-model', which is linear regression.

Bagging	Boosting	Other
Random Forest	XG Boost	Support Vector Machines
Decision Tree	Ada Boost	Linear Regression
Extra Trees		

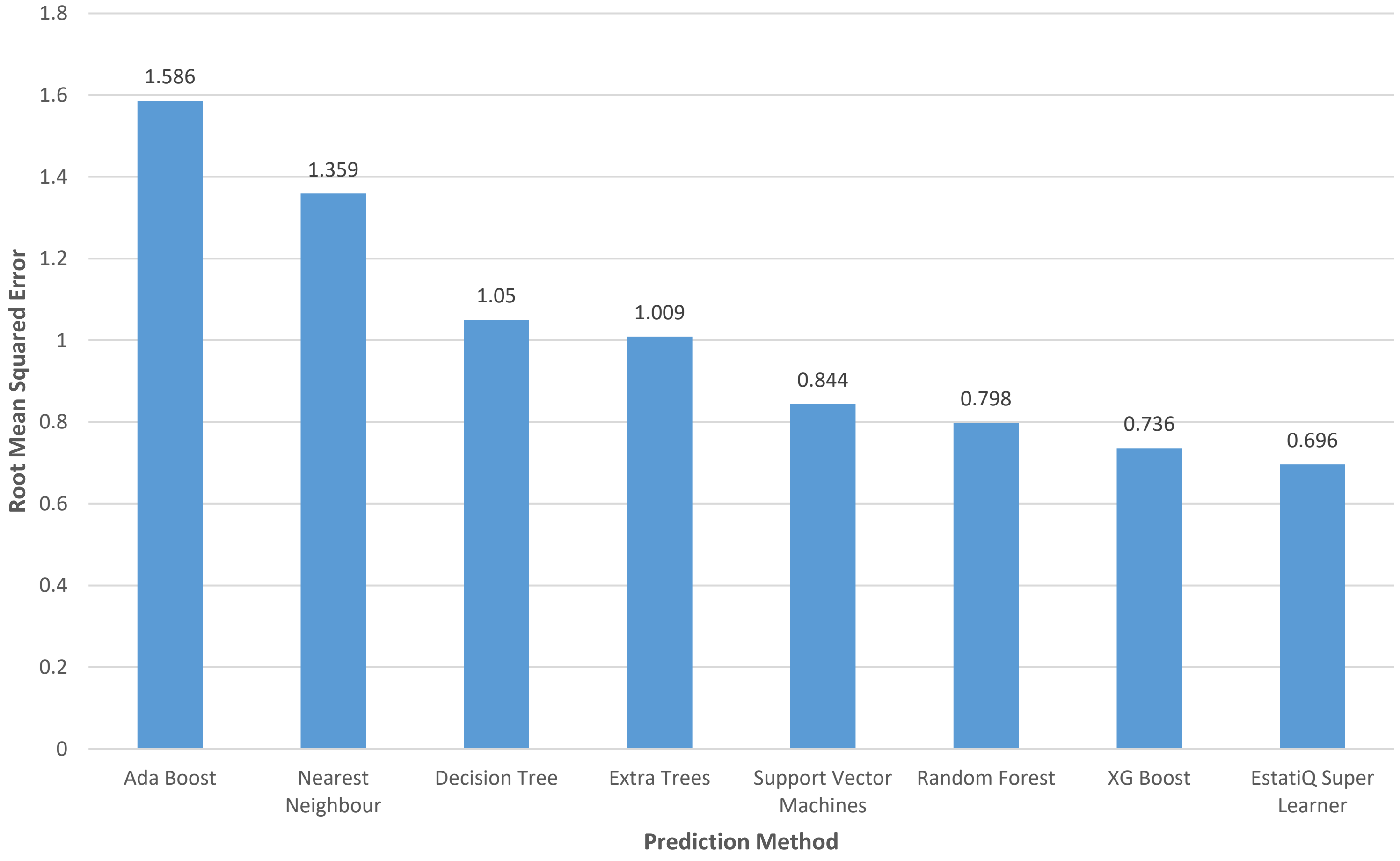


Results

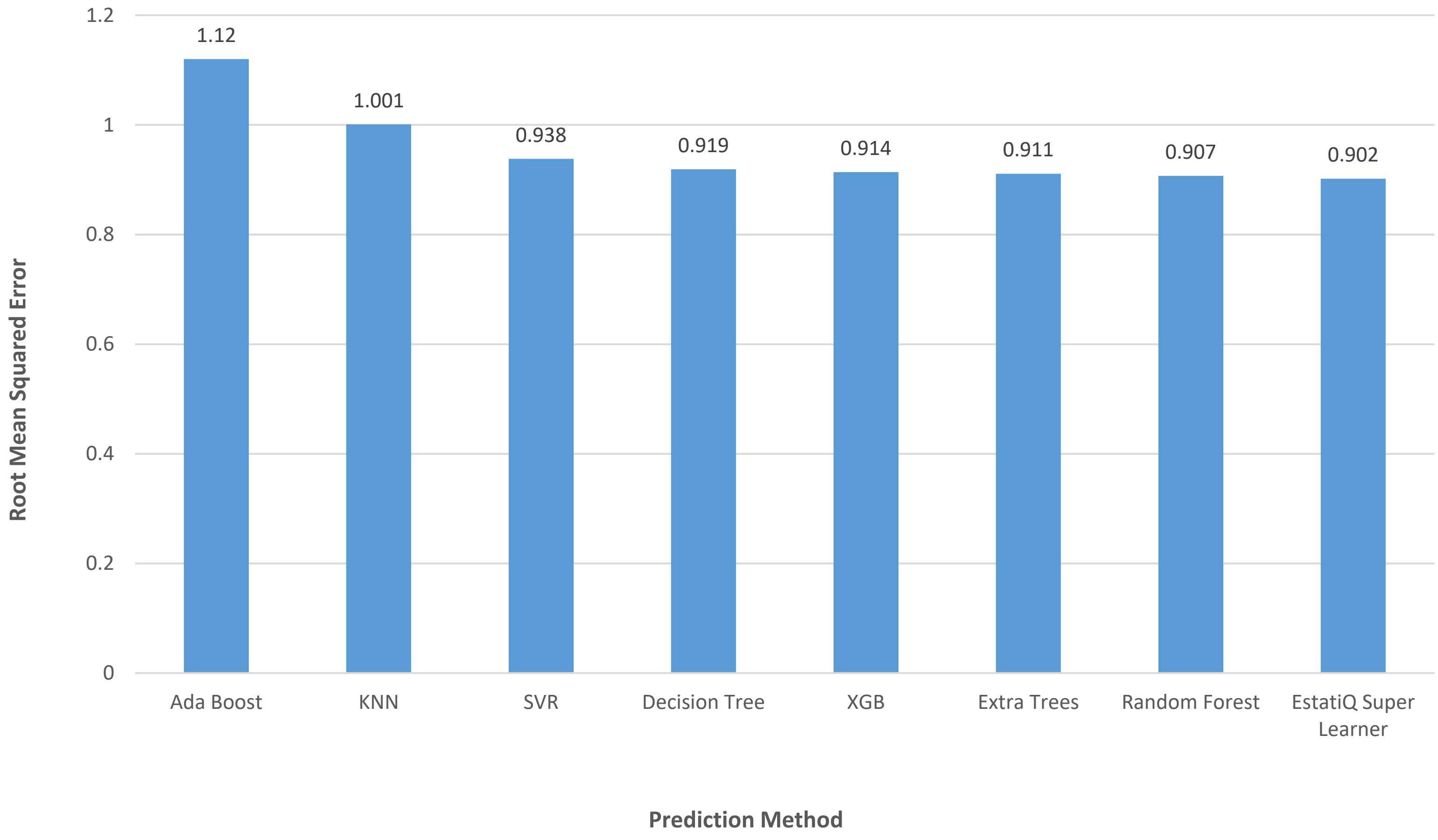
Buildings


1.2
1
0.8
0.6
0.4
0.2
0
Root Mean Sqaure Error
1.033
0.875
0.807
0.791
0.624
0.593
0.589
0.572
0.546
Ada Boost
Nearest Neighbour
Decision Tree
Linear Regression
Support Vector Machines
XG Boost
Random Forest
Extra Trees
EstatiQ Super Learner
Prediction Methods

Infrastructure



Land


1.2
1
0.8
0.6
0.4
0.2
0
Root Mean Squared Error
1.12
1.001
0.938
0.919
0.914
0.911
0.907
0.902
Ada Boost
KNN
SVR
Decision Tree
XGB
Extra Trees
Random Forest
EstatiQ Super Learner
Prediction Method

QINETIQ