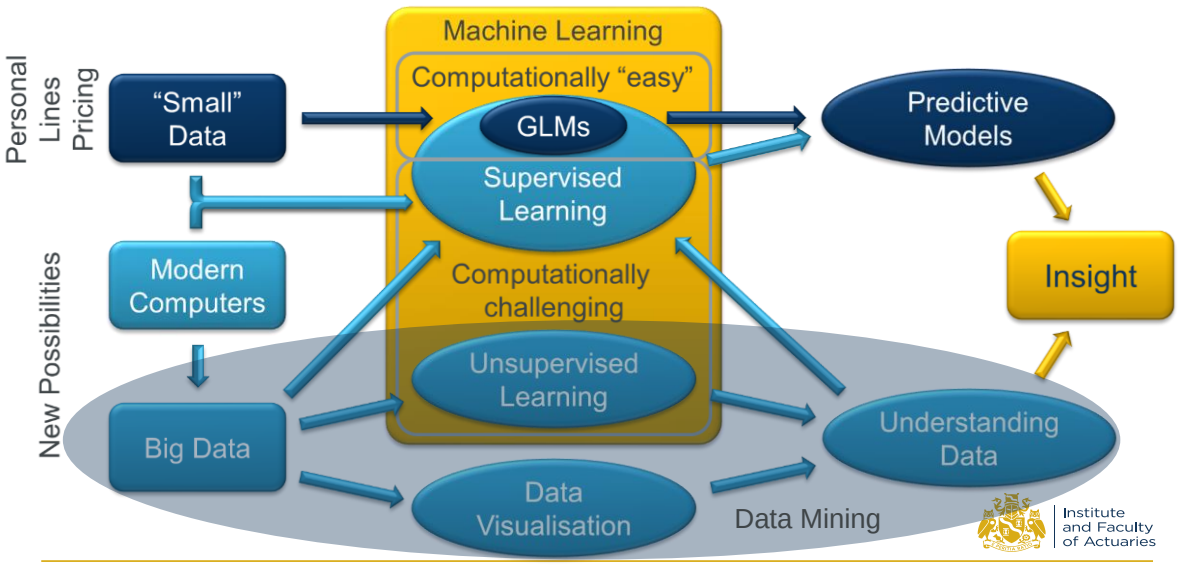
Institute and Faculty of Actuaries

# An Introduction to the Data Science Universe

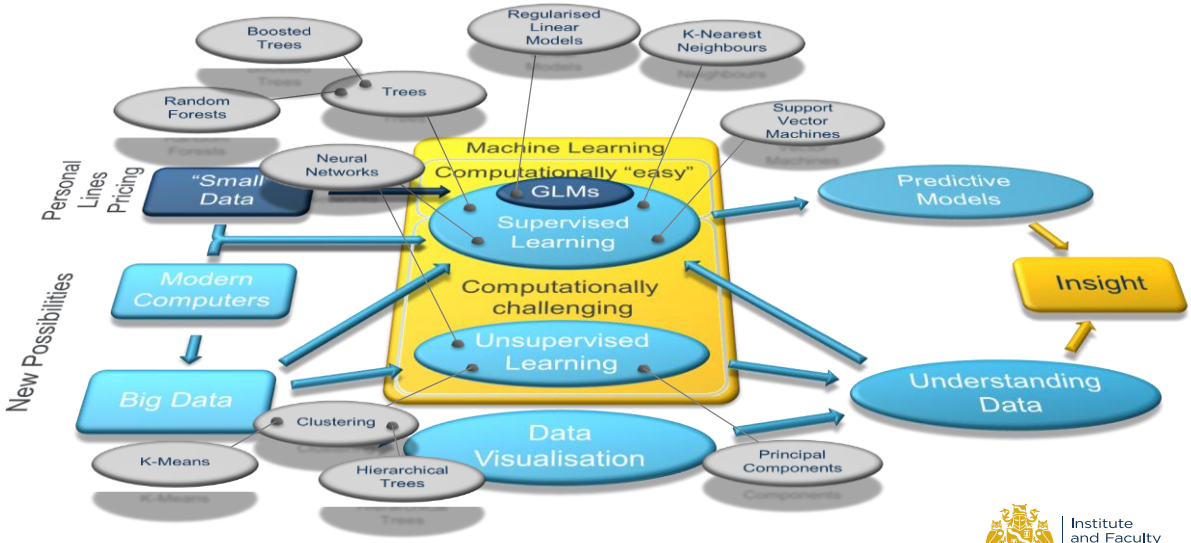
Mark Lee – Insight Risk Consulting
mark.lee@insightriskconsulting.co.uk

07 April 2016

## The Data Science Universe

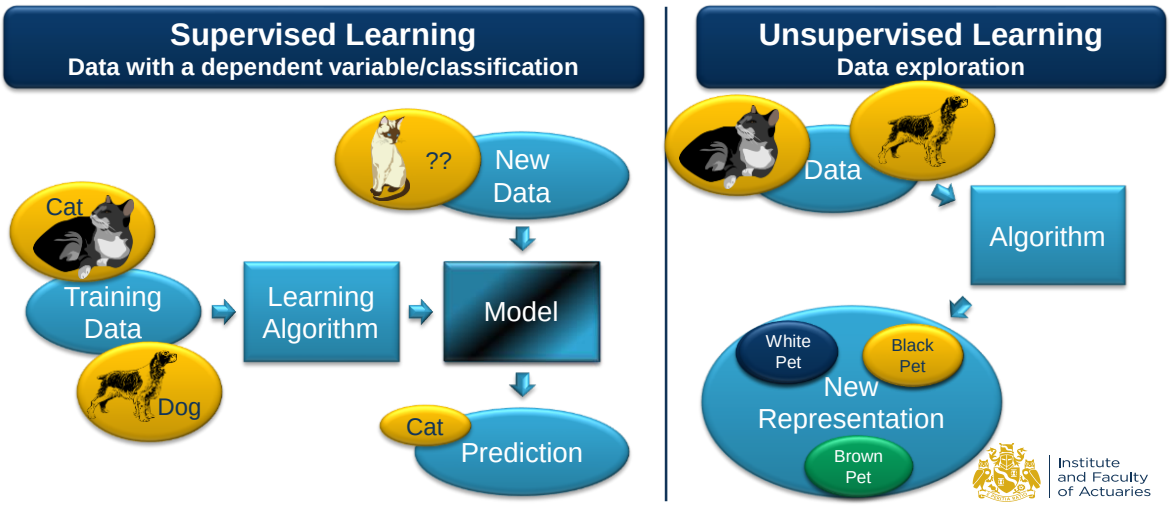


# The Data Science Universe

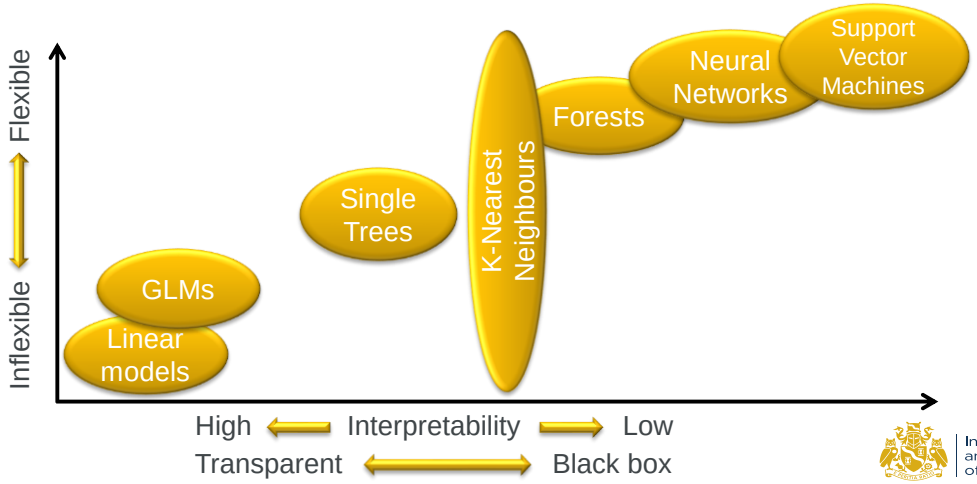





# Machine Learning

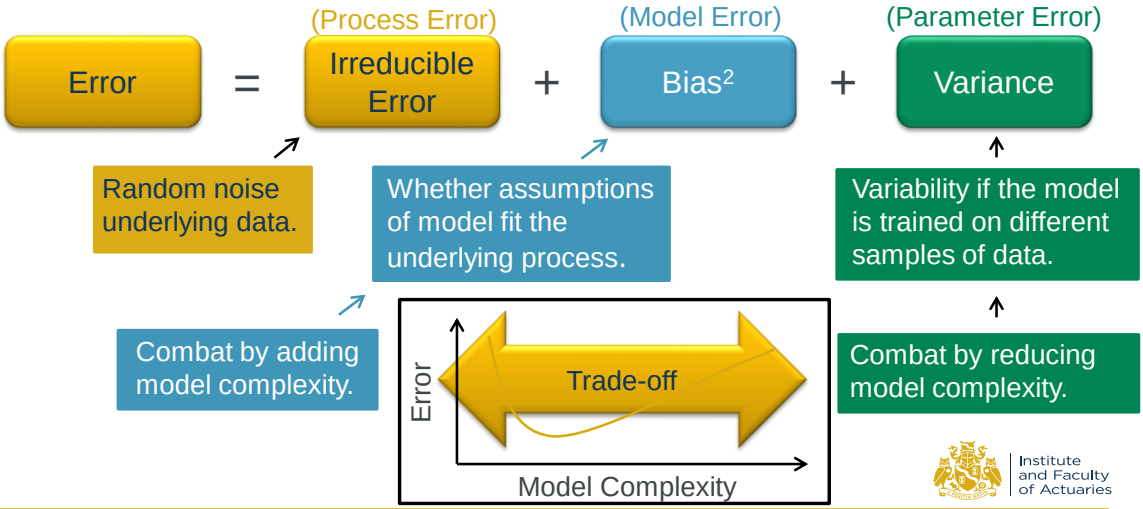


# Comparing Algorithms

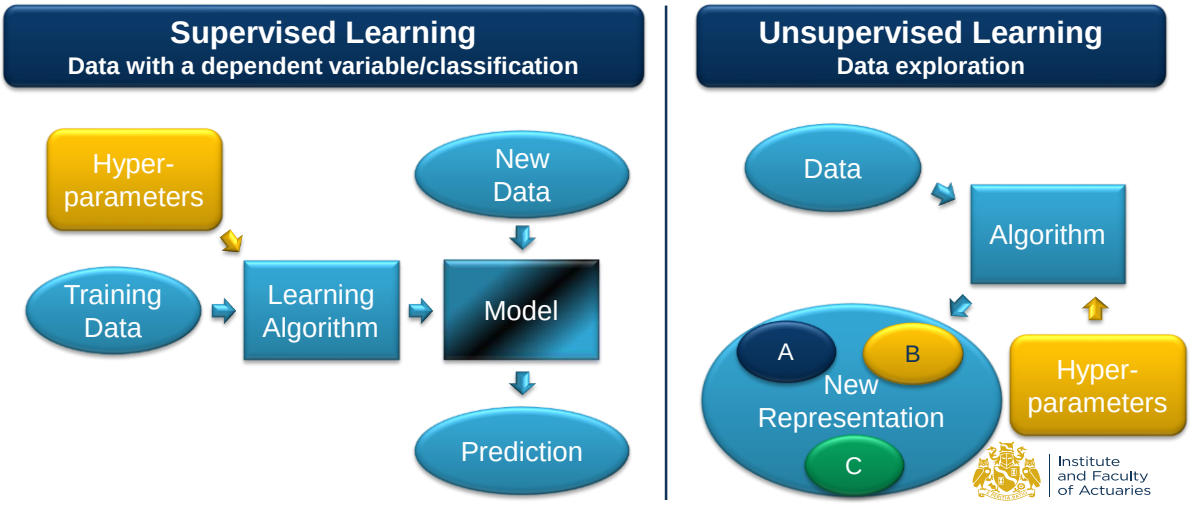





# Bias-Variance Trade-off



# Machine Learning






# So which algorithm is best?

- All algorithms are biased in some manner.
- The "**No Free Lunch Theorem**": no algorithm is best over all problems.
- Methods may work well for a certain class of problems.
- Best result may be an aggregate of methods.
- In other words, "**it depends**". Human judgement is required.

Institute and Faculty of Actuaries

## Issues for insurers

- New data sources – implications for better models:
  - Telematics for motor insurance.
  - Fraud detection via social media – several examples to date, but can it be automated?
  - Some health insurers are pricing based on health related data.
- More powerful algorithms:
  - Can GLM's keep up? Can new algorithms give a competitive advantage?
- How could London Market insurers gain from these techniques?

Institute and Faculty of Actuaries



## Issues for actuaries

- Will actuaries be replaced by machines/computer scientists?
- Human judgement is required - technically and statistically trained actuaries are well placed.
- Do actuaries need to train?
  - E.g., CAS Institute accreditation in data science.

Institute and Faculty of Actuaries

# References

- www.insightriskconsulting.co.uk
  - We mainly use R and Python platforms.
- No free lunch theorem:
  - D. Wolpert, *Neural Computation* **8**, p1341 (1996).
- Machine learning algorithms, book and video lectures:
  - *Introduction to Statistical Learning*, by James, Witten, Hastie and Tibshirani. www.statlearning.com
- LinkedIn group – search for “Data science for insurers group” on LinkedIn.com.