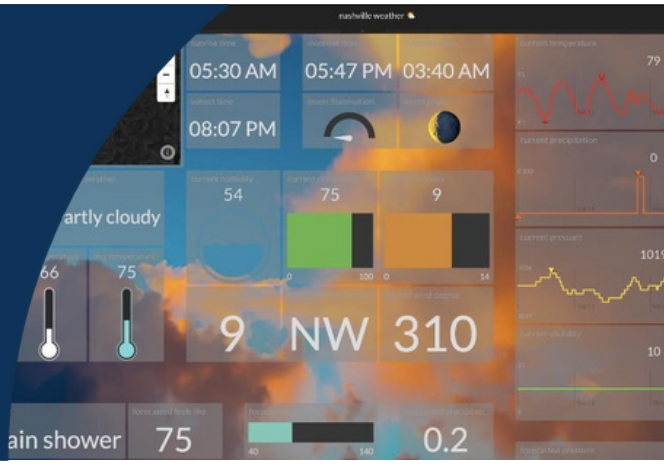
CASE STUDY

Weather Forecasting Organization

Data Science Training Academy



CLIENT CHALLENGE:

This European organization operates as research institute as well as a round-the-clock service producing global numerical weather predictions and other data for its member states and broader community. Their Scientists, Statisticians and Programmers need to manage a high volume of rapidly changing data on a constant 24x7 basis. To achieve this, these roles need to be on the forefront of programming skills to leverage technology for accurate forecasting.
Leveraging our model of using practising Consultants to deliver training, Learning Tree was invited as a trusted training partner to support this business need through developing a Data Science Academy addressing the following upskilling:

- Jupyter notebooks
- Supervised and unsupervised learning
- Deep learning and interpretability
- Data mining and pattern discovery
- Mathematical basis of machine learning
- Uncertainty quantification
- Gaussian processes

The training was required to be 3-4 days of consistent training for Statisticians and Programmers.

OUR SOLUTION:

Through engagement and discussion with our Lead Data Science Consultant, a two-tiered customised solution was firstly developed as a specification and then on approval as a programme of courses.

The first course 'Data Science I' addressed the following learning outcomes from the perspective of a Python programmer:

- Understand the fundamental ideas behind artificial intelligence and machine learning in particular.
- Choose algorithms that would be good candidates to solve a given problem.
- Interpret the measurements used to gauge the quality of a model.
- Appreciate the workflow needed to produce machine learning solutions.

The next tier course 'Data Science II' was designed with the following learning outcomes:

- Understand the fundamental ideas behind
- deep learning
- Generalise machine learning concepts to
- Bayesian analysis and Gaussian processes in
- particular
- Interpret the measurements used to gauge
- the quality of a model

This programme is unique in its emphasis on a Bayesian approach to modelling rather than the more commonly used "frequentist" approach. The Bayesian approach allows one to combine prior information you have about the phenomenon you are trying to predict, with evidence from newer information to make a better prediction. This approach better fits the way a weather prediction organization continuously gets new data. Until the recent past, the Bayesian approach has not been practical because only a few special cases can be solved by hand mathematically. With current computational resources and Monte Carlo methods implemented in Python, it is now a feasible approach to prediction.

Continued on next page

The Results

95%
LEARNER SATISFACTION In the course

4.0 ★★★★
AVERAGE SCORE of course instructors

3.77 ★★★★
AVERAGE course score

To date, Data Science 1 has been delivered three times, upskilling 43 Scientists, Statisticians and Programmers. Learner feedback was exceptionally positive with high feedback scores using the Kirkpatrick evaluation model.

- "A very good overview of the different data science methods, the Jupyter notebooks were really helpful."
- "Very good and useful course with lots of theory and practical examples."
- "Christopher was an excellent lecturer, very clearly explaining everything."
- "Was very knowledgeable and provided immediate assistance with any questions."
- "A great effort was made in explaining the theory behind each concept, at the right level of detail."
- "The Trainer made the course very attractive, and he was well organised with the time allocated between teaching and exercises and kept good time schedule. The trainer was helpful when we were in trouble, always available to answer our questions and to explain the outcome of the exercises. The course was very well organised, everything technically worked properly. Very good experience!"

Data Science II has been delivered once upskilling 15 Scientists, Statisticians and Programmers. Again, feedback has been very positive.

- "Very good course programme! We learned so much during these 3 days! I really enjoyed the practical exercises, helped me to understand better these concepts."
- "Very well presented and very clear explanations using practical examples and nice exercises. You cover so much in only 3 days. I really enjoyed every minute of it! Now I am ready to apply some of those ML concepts, very exciting!"
- "Very nice and competent instructor. Great ability to transmit knowledge. Thank you."
- "Very professional teacher, overall, he provides a great learning experience."
- "Thanks for this time thanks for this mind opening experience."
- "Thanks to Mr Mawata clarity and expertise for this course."

Based on this positive start to the Academy a further five deliveries were scheduled in 2021 as the academy expanded to a wider set of roles. As part of our continuous improvement cycle, we look to see how each course can be improved to increase the benefits to the learners and their organisation. The Tier II course has been broken out into two shorter modules with one devoted to Bayesian Data Analysis (e.g., Bayesian estimators/regression, MCMC, Gaussian Processes) and another course solely dedicated to neural networks (and possibly a bit of decision trees/random forests).