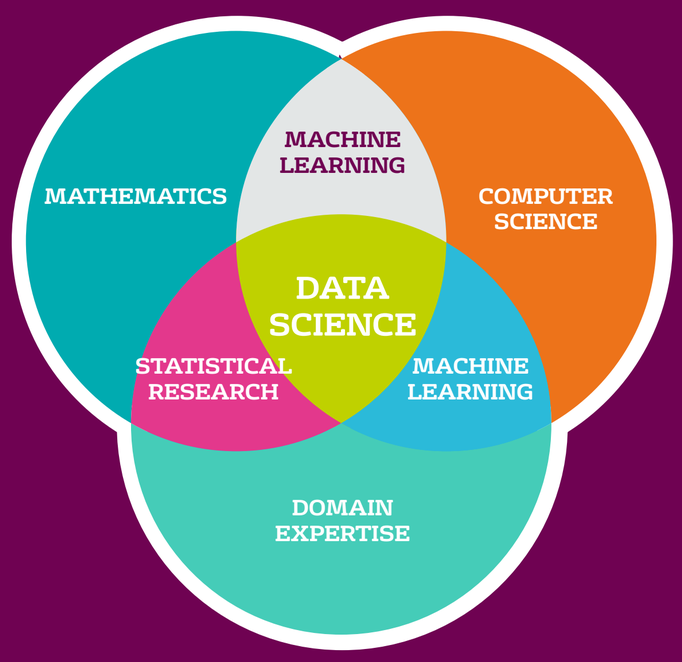
What is Data Science & AI?

Data Science

Definition

- One definition (of many) many): Extracting knowledge from data a nd applying it to problems in a wide range of fields.
- Skills : Computer science, statistics, information science, mathematics, information visualization, data integration, graphic design, complex systems, communication and business….and more!
- Applications : Anywhere there is a gap in understanding and enough data to fill it!



AI

Definition

- Classical AI aims to teach machine deduction for specific tasks: learn rules from data…like we do when learning maths or a new language!
- General AI attempts to create general inteligence that can learn to perform any task that human intelligence can attempt.
- Often, but not required to, mimic human reasoning
- Sometimes it is not a good idea to use AI! Ethics must always come first. Ask yourself the question: is my AI going to cause more damage than good?
- AI fed with biased data will have those same biases. Trash in, trash out

A day in the life of a data scientist

Problem familiarisation

- Make aims clear
- Background information
- Speaking to domain experts

avoid disconnect between data scientist and domain expert in order not to reinvent the wheel

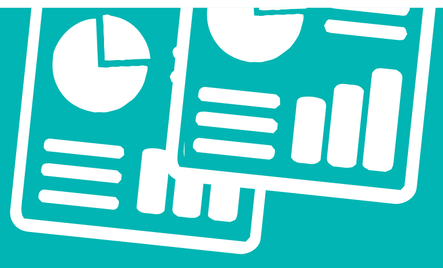
Data cleaning

Data comes in all shapes and forms: numbers, images, videos, text, audio files, etc. Most of a data scientist's work involves some sort of data manipulation and normalisation such as

- Formatting
- Aggregation
- Sorting
- Collect extra data if needed

Initial testing & exploration

- Plotting / data visualization
- Summary statistics
- Finding the features that best describe the dataset
- Sufficiency and quality of data



Reflection & reporting

- Data visualization
- Statistical findings
- Conclusions of study

Testing & validatation

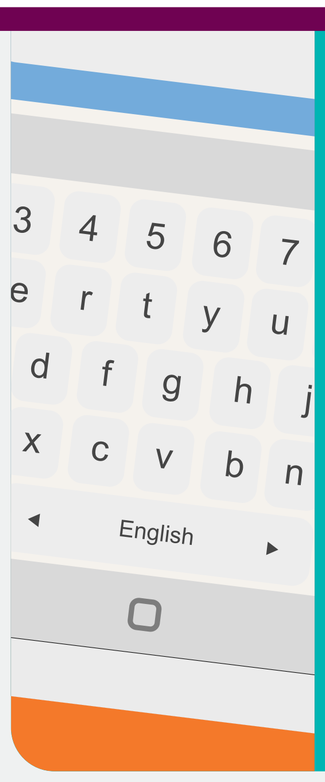
- Making and testing predictions using your model
- Testing approach on unseen data

Analysis & modeling

- Using models to describe the dataset (both Machine Learning and traditional methods)
- Coding & Debugging
- Quantify uncertainties
- Sufficiency and quality of data

Who can become a data scientist

- Everyone! (If you enjoy statistics and coding).
- The data science community is expanding and skills are needed from almost every domain.
- You do not need formal training in STEM to become a good data scientist.
- Plenty of resources are available online to dip your toes into the field and to start engaging with the community!



Applications

- Have you ever noticed how your phone's keyboard can predict your words?
- How about e commerce websites magical ability to always know what you want to purchase?
- On a day to day basis, we are involved with a myriad of data science fuelled applications, such as voice assistants (Siri, Alexa, Google Now), spam filtering, google autocomplete functionalities and so on and so forth
- Ongoing research in a huge range of fields (diagnosing diseases, saving energy, monitoring structural health of buildings, self driving cars, virtual reality, creating art….)

University of BRISTOL
Jean Golding Institute

bristol.ac.uk/golding

@JGIBristol