**Discussion Topics: Behavioural Data Science**

*Disclaimer: I'm not an expert in Big Data methods, so some of the answers are based on my general understanding of it, do take them with a pinch of salt.*

(1) Is there any advantage in using large datasets (Big Data) for understanding human behaviour? Support your opinion with examples/evidence.

**Some Advantages**

- More accurate prediction of behaviour usually due to large *volume* and *veracity*?
- Data collected is realistic compared to say lab experiments: analysis has higher external validity (more practical in nature). Just like in field experiments, there is (possibly) a richer set of contexts and variables which can possibly help enrichen our perspectives?
- If we can open the black box, may find correlations between variables and predictions which one may not usually expect: Inspires new theory and experiments to explain this?

**Some Disadvantages**

- Depending on the kind of analysis used, may be a black box: don't really understand what is going on: the more data there is, the more complex the interactions: harder to understand the effects? I.e. Accurate prediction not the same as good understanding.
- Data while big, is exhaustive, usage of proxies: Correlations if found are still subject to omitted variable bias, reverse causality: trying to use exhibited correlations between inputs and predictions to understand WHY people make certain choices might be pointless? Goodhart's Law?
- For certain machine learning models, data is trained to a specific sample set: may fail when confronted with unfamiliar patterns: e.g. self-driving vehicles, further models may not be able identify when it is unknown. System is not aware of its limitations.

(2) What data types exist and how different data types impact on the value/worth of personal data?

- Traditional Data: Data which is already being collected (naturally) alongside people's standard activities: e.g. water/ electricity expenditure in utility bills, companies expenditure data etc. Only possibly influences behaviour in retrospect.
- Invasive Data: Data which is collected intentionally and sometimes provided alongside other products, may affect people's behaviour. (e.g. Fitbit, Survey apps?)
- Inventive Data: Content vs Metadata: Metadata gives additional context about the main issues of interest (Involves coming up with new ways of collecting additional supporting data; something like "control variables").

These are usually increasing (down the row) in the costs of providing the data; e.g. to the consumer, one loses more and more privacy and hence data is worth more to them. On the other hand, the value of the data to the purchaser typically also increases (down the row).

Reasons for increasing value to the demander:

- Aggregation, increasing returns to scale: data is only useful if purchased together and not one by one.
- Increasing dimensions of the data mean more returns out of the data. In inventive data, contexts help to raise the marginal benefit from data from a single consumer.

Ganna mentions how the trade of data at these different levels is affected by demanders' perceptions of the data. At traditional data (and also invasive data) there are issues with the quality of the data which prevents direct trade: i.e. there is a wedge between 3rd party aggregated data and data collected from an individual (though there may be no actual difference).

(3) What is a sentiment analysis? How can sentiment analysis be used to better understand human behaviour? Provide examples.

Sentiment analysis refers to analysing the emotional content of various forms of data (e.g. text). In psychology/ neuroscience, they often talk about people having hot/cold states of decision making, the former being more emotional while the latter being more rational. Emotions can adoption of reference points and learning via conditioning and may further interact with people's rational behaviour (?); e.g. storage and recall of memories may be conditioned by emotions. Furthermore, empathising with other's emotions might also lead to shifts in own emotions and behaviour as well. Hence, it seems that by incorporating emotions as new variables within data might be important for understanding human behaviour.

Application wise, Ganna talked about how they used the text/conversations over the span in movies to measure the dynamics of the emotional content (+ve/-ve). This is then associated with the popularity of the movie.
Other possible contexts: political decisions, stock market decisions, emotional contagion (using news, speeches, videos, social network data?) etc.

(4) Without going into too much technical detail, explain how standard machine learning models are different from deep learning models.
E.g. for image recognition:

- Standard machine learning: single layer, less data needed. Programmer needs to define the relevant dimensions in the data. Fast, but less accurate.
- Deep learning: multi-layer, needs more data. Algorithm can learn to identify what are the relevant dimensions in the data. These are like latent variables associated with the original set of data. Slow, but more accurate.

Both of them are black box in the sense that they make predictions given a certain set of data, but it is not clear what is the direct effect of given inputs: this could be because there are too many complex interactions for there to be a clear idea (even if the algorithm used is clear). (Not like in a regression table).

**Reference for what is meant by black box:**
https://towardsdatascience.com/the-black-box-metaphor-in-machine-learning-4e57a3a1d2b0

(5) What is "explanability"? Why is "explanability" important when we think of the artificial intelligence (AI) modelling?

Explanability refers to whether one is able to tell what (and why) the algorithm is doing something at each step. Ethical issues are related to this issue of explanability: If we do not know why and how the system is doing something, should we be allowed to use it? For example, will ML systems by learning from biased data, also exhibit (even greater) discriminatory behaviour? Are they able to make decisions based on learned criteria which do not impinge on our moral values?
A related problem here is that these values are not learned by AI systems (yet): so if we can't be sure that they they are moral, can we trust their decisions? Hence, explanability is important as it gives humans a chance to second guess the AI's decisions.

Perhaps they may be more applicable in situations with less ethical risk and with less direct intervention on people's lives: e.g. which area is high security risk and needs to be patrolled more.

**Sources:**
https://www.siliconrepublic.com/machines/ethics-deep-learning-4irc
https://towardsdatascience.com/7-short-term-ai-ethics-questions-32791956a6ad
https://arxiv.org/abs/1606.06565.pdf

**Increasing explanability using anthropomorphic machine learning?**

From my understanding, anthropomorphic machine learning sort of adapts some aspects of decision theory to machine learning, ways in which this could be done:

- Applying shortcuts from predictions from decision theory (which are established): e.g. can save on "learning" computation in the prediction from latent attributes?
- The psycho-neurological perspectives of how people actually learn things into designing AI learning processes. Link below mentions how it possibly allows for a better understanding of how the AI is making its decisions (see link below)

Additional Information on Anthropomorphic machine learning:
http://eprints.lancs.ac.uk/126498/1/CO_COMSI_2018_05_0087.R1_Gu.pdf