Data, AI and Decision Making

Distinguished Professor Jie Lu
Fellow of IEEE
Director of Centre for Artificial Intelligence
University of Technology Sydney
Jie.Lu@uts.edu.au
Main AI Technologies

Computer Vision
- Image Recognition
- Object tracking

Machine Learning
- Deep Learning
- Transfer Learning
- Predictive Analysis

Computational Intelligence
- Fuzzy Systems
- Neural Network
- Evolutionary Computing

Natural Language Processing
- Translation
- Information Extraction

Logic and Expert Systems
- Reasoning
- Knowledge base

Planning, Scheduling & Optimization

Robotics

Brain-Computer Interface
Machine learning provides opportunity to understand data and interpret it.

- How to support decision making by applying machine learning on data?

Big Data Era
- High-Speed
- Data Streams
- Huge Volume
- Multiple- Sources

Algorithms Models

...
Data:

Structured vs. Unstructured Data

- Multimedia data (images)
- Graph data (social network)
- Streaming Data (IoT)
**Static data**: Data is prepared in advance and stored in databases. The distribution variation between training data and new data is usually ignored (or treated as overfitting problem).

**Data stream**: Data is received in stream. Unpredictable variation of data distribution may occur. There are two types of machine learning models for streaming data.

---

[Diagram showing the process of training and predicting with static data and data stream.]
Machine Learning

Example of Machine Learning:
- Transfer Learning
- Deep learning
- Concept drift learning

Machine learning is a computational process of discovering patterns from historical data.

The goal of machine learning is to extract useful information from data and transform it into an understandable structure (knowledge) for further use.
User interest prediction: not enough labelled data in the book dataset however, a lot of labelled data is available in the movie dataset.

The prediction model of Movies cannot be directly used to users’ prediction of Books.

1) Transfer Learning Solves Labelled Data Insufficiency
2) Concept Drift Learning on Data Stream

- Concept drift at time $t$ can be defined as the change of joint probability of $X$ and $y$ at time $t$, where: $\exists t: P_t(X, y) \neq P_{t+1}(X, y)$

- It is the root cause of accuracy degradation of evolving data stream learning, which leads to the increasing of learner error rates.

Distribution drift causes accuracy degradation or error rates increasing

When is the best time to update a learning model?
AI supports decision on customer services

--- Personalised Recommender Systems

User-item Preferences Data (Ratings, Purchase,...)

Recommender System

Predict items that may be interested to a user

Type:
- Collaborative filtering
- Content-based
- Others

User-Items matrix

New Developments:
- Tree RS
- Group RS
- Cross-domain RS
- Interest draft RS

User1: 3 5 5 1 ?
User2: 3 5 4 1 5
User3: 1 3 3 4 3
User4: 2 2 2 2 2
User5: 3 1 2 2 1

AI supports transportation management

**Bus replacement background**

- Sydney Trains provide Trackwork Transport Services for passengers when the use of rail network is disrupted.
- Sydney Trains spent approximately $35 million on replacement buses in FY 2015/16, but still facing travel delays and customers dissatisfaction.
- To develop AI technology to optimise customer experience whilst managing the costs of the replacement services.

**Data driven approach**

- Opal Tap-on and Tap-off data, Train OTR data
- Google historical travel time information from RMS
- Operational constraint knowledge from in-house bussing team
- Machine learning to find hidden patterns and give insights of the data

**AI model produces better bus timetable**

<table>
<thead>
<tr>
<th>Replacement Run</th>
<th>Enoggera to Chatswood</th>
<th>Wrenya to Hornsby</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Timetable</strong></td>
<td>Real Timetable</td>
<td>Model</td>
</tr>
<tr>
<td>Bus Hours</td>
<td>119</td>
<td>119</td>
</tr>
<tr>
<td>Average Journey Time mins</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>Average Bus Waiting Time mins</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Average Train Waiting Time mins</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Occupancy rate</td>
<td>53%</td>
<td>53%</td>
</tr>
</tbody>
</table>

- Within the same running cost, AI model leads to shorter waiting time and travel time.
- Constrained by the same travel time, AI model leads to significantly reduced running cost.
Challenges in AI applications: Switch type of our jobs

- If the AI doctor fixes 100 people and makes mistakes on two, for example, and the human doctor treats 100 people and makes five mistakes, which one is better?
- **AI will never be perfect. It just needs to be better than people.**
- If machines are quicker and cheaper than humans, if people make more mistakes, let’s deploy AI?

The Impact of AI –

*It may require million people to switch job categories entirely.*
AI provides a great opportunity to improve our wellbeing and lift the economy through data-driven decision support.