VICTORIA UNIVERSITY OF WELLINGTON
Te Whare Wananga o te Upoko o te Ika a Maui

School of Engineering and Computer Science

COMP 307/AIML420 — Week 12

Lectures 23-24

Other AI Topics and Other Information

Mengjie Zhang

Mengjie.zhang@ecs.vuw.ac.nz

Outline

- Other Topics
  - Support Vector Machines
  - Knowledge based systems
  - Natural language processing/Text Mining
  - Data mining and knowledge discovery
  - Big data and Data Science
  - Deep learning
  - Transfer Learning
- Other Information
  - New AI Courses (brand new)
  - New Postgraduate AI qualifications
  - New BSc AI Major (brand new)
  - Scholarships for Summer and Honours at VUW
  - AI Research Scholarships for Summer, Hons, Masters, PhD
  - Potential AI Projects for Summer and Honours
  - VUW AI/EC Research Strengths and International Leadership

Topics Covered So Far

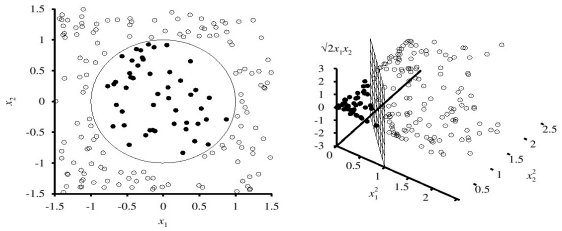
- AI establishment/philosophy/history
- Search
- Machine learning
  - basic techniques and concepts
  - neural networks and learning
  - evolutionary computation and learning
- Reasoning under uncertainty
- Bayesian networks
- Planning and Scheduling
  - Routing

- Many other AI topics that have not been covered

Support Vector Machines

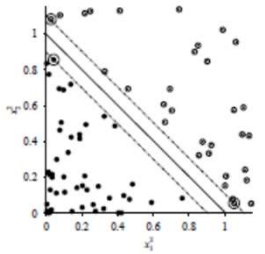
- One of the most powerful classification methods
- Can also be used for regression
- Primarily developed for binary classification
- Have good ideas/properties
  - Efficient learning (like **simple perceptron**)
  - As non-linear as they need to be (like **big NNs**)
  - Largely immune to overfitting (not always)
- Based on Structural Rick Minimisation
- Use clever Maths …
  - Learning theory (PAC Learning)
  - VC dimension

Support Vector Machines (2)



- Non-separable in 2D but Separable projected into 3D
- "everything is separable, in enough dimensions"
- SVMs can effectively project data into very high dimensions
- Classify by learning a "perceptron" "up there"
- Use clever maths makes this possible without actually carrying out the projection!

Support Vector Machines (3)



- Only some of the training examples are needed to define the boundary
- Known as *support vectors*
- #support vectors << # training examples

Support Vector Machines – Case Study (4)

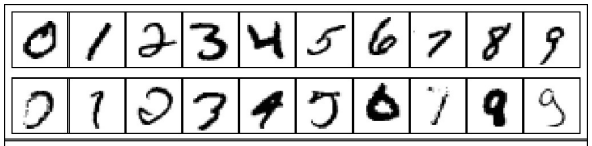


Figure 20.29 Examples from the NIST database of handwritten digits. Top row: examples of digits 0–9 that are easy to identify. Bottom row: more difficult examples of the same digits.

- 3-NN: 2.4% error
- 400-300-10 NNs: 1.6% error
- LeNet 768-192-30-10: Network: 0.9% error
- LeCun shared weight/convolutional network: 0.7% error
- SVMs: 0.6% error (1995)
- DL? …

Knowledge Based Systems (1)

- KBS (technology term) = Expert Systems (task based term)
- Expert systems = Knowledge Engineering = (Symbolic)AI
- KBS = AI …

- An Expert system is "an intelligent computer program that uses knowledge and inference procedures to solve problems that are difficult enough to require significant human expertise for their solution. The knowledge of an expert system consists of facts and heuristics" [Feigenbaum]
- Knowledge Representation:
  - Rules, DT, semantic network, Frames, Scripts, OORepresentation

Knowledge/Rule Based Systems (2)

- A powerful knowledge representation scheme is

  "if – then" rules:

  if *<description of situation>*
  then *<consequence>*

- Basis of "rule-based systems" or "expert systems"
- First big success story of AI
  - AI systems that solved real problems that couldn't be solved by traditional programming. E.g. R1 for configuring Vaxes
  - Captured Experts' knowledge in if-then rules
  - Reasoned with the rules to do diagnosis, classification, design…
  - Could even explain their reasoning
  - but over-hyped, collapsed,
  - emerged again as "Business Rule systems", …

Knowledge Based Systems (3)

- A typical ES architecture

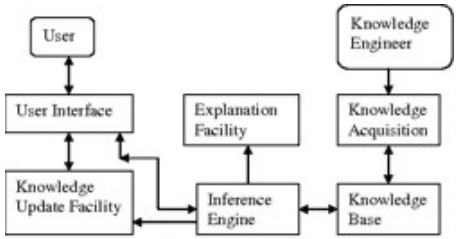


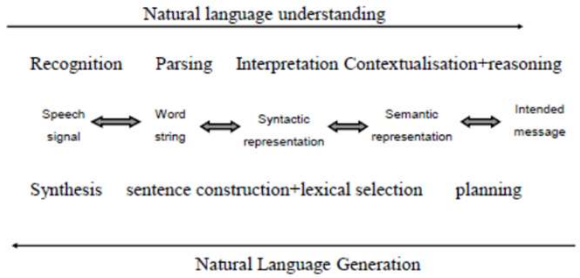
Figure 1: Architecture of Expert System

Knowledge Based Systems (4)

- Examples of Knowledge-Based Systems
  - MYCIN: blood infection diagnosis
  - PROSPECTOR: analysing geological data
  - XCON/R1: configuring computer orders
  - GASOIL: designing gas-oil separation systems for offshore oil platforms
  - Mortgage approval
  - Building code checking
  - Detecting credit card fraud
  - Plastic formulation
  - Aircraft wing design
  - Copier paper-transport design
  - Help desk (Compaq)
  - Printer configuration aid (MS Windows)
  - …

Natural Language Processing (1)

- One of the oldest topic
- Also a hottest topic in AI



Natural Language Processing (2)

- History
  - In the early 1960's there was great optimism about the ability to perform machine translation (MT). This was of particular interest to the military in Western Europe and the US.
  - This initial optimism was dashed when it was realized that NLP is full of hard problems.
  - The situation nowadays is rather better, with a number of (commercially) successful NLP systems on the market.
- Application areas:
  - Machine translation, Speech-to-speech translation
  - Text/Web mining, Information extraction, Information retrieval
  - Database querying, Instruction following
  - Authoring aids (word processing tools)
  - Text, speech, images, robots, social media, …

Natural Language Processing – Stages (3)

- Morphological processing: find root form and category, eg., "rose": past tense of "rise" or singular noun.
  - Involves lexical and morphological knowledge (how to make plurals, past tense, present continuous tense, adverbs from adjectives, etc).
- Syntactic processing: parse sequence of words into constituent structure.
  - Involves lexical and syntactic knowledge: grammar, number and case agreement, etc.
  - Syntactic constraints may resolve some lexical or syntactic ambiguities.

Natural Language Processing – Stages (4)

- Semantic processing: assign meanings to syntactic structures. Constraints about entities and actions are applied.
- Discourse processing: interpret sentence in context of discourse. Entities referred to must have been introduced, and overall discourse structure must be consistent. Can help to resolve ambiguities.
- Pragmatic analysis: reinterpret structure representing what was said to determine what was meant. Can use a wide range of knowledge about the world, and in particular, knowledge about peoples' beliefs and goals and conversational conventions.

Data Mining and Knowledge Discovery(1)

- Why data mining?
  - Data comes like water out of a fire hydrant. You can't drink it (Anon).
  - We are drowning in information but starving for knowledge (John Naisbett).
  - Hardware advances in data collection and storage have far outpaced
  - software advances in data analysis and manipulation.
  - Organizations collect more data than they can handle.
  - Data that may never be analysed is still collected out of fear of
  - missing something that might be important.
  - As databases grow, decision making directly from their contents
  - is not feasible; knowledge derived from the data is needed.
  - Supermarket chains, credit card companies, banks routinely
  - generate daily volumes of 10-100GB -→ ???
  - Scientific and remote sensing instruments collect gigabytes of
  - data everyday.

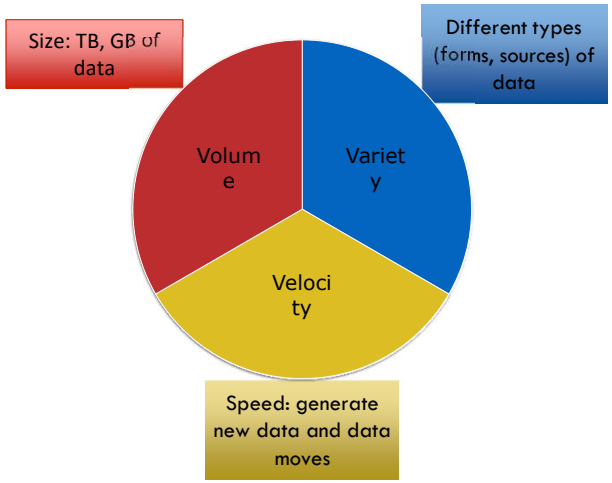
Data Mining and Knowledge Discovery (2)

- Knowledge Discovery in Databases (KDD) is the non-trivial process of identifying valid, novel, potentially useful and ultimately understandable patterns in data [Fayyad].
- The non-trivial extraction of implicit, previously unknown and potentially useful knowledge from data [Adrians]
- KDD/DM is not a new technique but rather a *multi-disciplinary* field of research: all make a contribution
- Examples
  - Fraudulent credit card transactions, Good/bad loan risks
  - New class of stars
  - Put beer and disposable nappies together and you may sell more of each
  - Put perfume and greeting cards together and you'll sell more of each
  - Inspect credit card transactions, find people who brought scuba
  - gear and lessons and send discount coupons for Carribean cruise
  - Recognition of market segments that respond to specific characteristics
  - Ineffective advertising
  - Recognition of a particular face in a database of photographs
  - Finding all cyclones in a database of satellite images

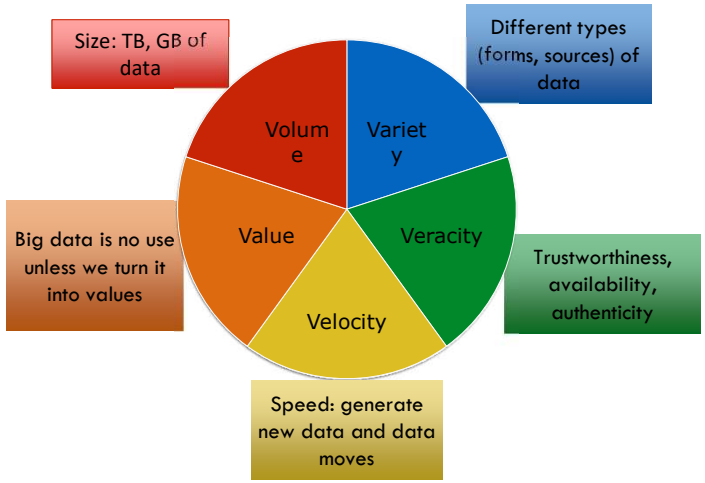
Big Data (1)

- Big data is a term for data sets that are so large or complex that traditional data processing applications are inadequate
- Characteristics
  - Volume: The quantity of generated and stored data.
  - Variety: The type and nature of the data.
  - Velocity: The speed at which the data is generated and processed to meet the demands and challenges
  - Variability: Inconsistency of the data set
  - Veracity: The quality of captured data can vary greatly

Big Data – 3Vs (2)



Big Data – 5Vs (3)



Big Data – Big Dimensionality (4)

- Thousands of features
- Millions of features
- Not many of them are useful
- Some are irrelevant/useless
- Some are redundant
- Features are not independent but correlated/interacting
- Useful features are not equally important
  - Automatics feature selection
  - Automatic feature construction/extraction
  - Big dimensionality reduction

Big Data -- Applications (5)

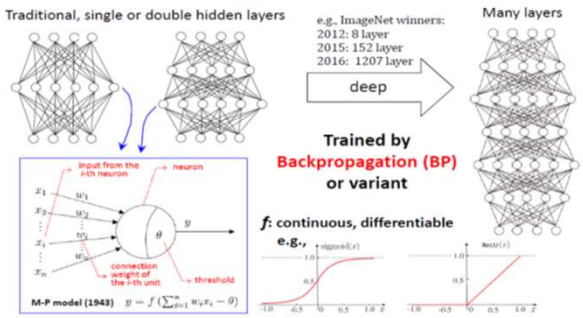
- Big Companies: Google, Software AG, Oracle Corporation, IBM, Microsoft, SAP, EMC, HP, Dell, etc.
- Governments
- Manufacturing
- Healthcare
- Education
- Media
- Social networks
- Science and Engineering
- Sports
- Research

Deep Learning (1)

- A hot topic – partially due to the recent AlphaGo (Google) event
- No uniform definitions, but several key points:
  - multiple layers of nonlinear processing units
  - supervised or unsupervised learning of feature representations and transformation in each layer
  - with the layers forming a hierarchy from low-level to high-level features
  - Sufficient complex models for processing hard problems
  - Requiring a large number of training instances
  - Based on distributed representations and processing e.g. GPUs
- Typical supervised learning algorithms
  - Logistic Regression
  - Multilayer perceptron
  - Deep Convolutional Neural Networks

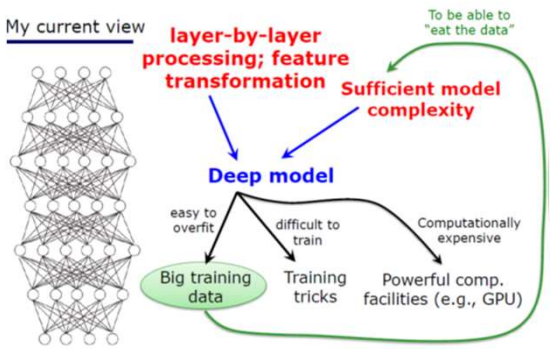
Deep Learning (2)

- What is "Deep Learning"? Deep Learning = Deep Neural Networks?



24

Deep Learning [Zhou] (3)

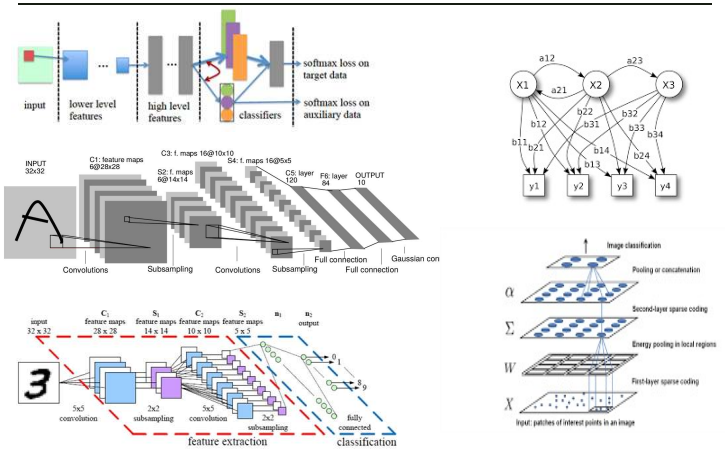


25

Deep Learning (4)

- Typical unsupervised learning algorithms
  - Auto Encoders, Denoising Autoencoders
  - Stacked Denoising Auto-Encoders
  - Restricted Boltzmann Machines
  - Deep Belief Networks
  - LSTM, GANs …
- Non-NN Deep learning algorithms
  - PCA-Net, Deep Forest, GP, …
- Applications
  - Image recognition
  - Automatic speech recognition
  - Natural language processing/text mining
  - Drug discovery and toxicology
  - Recommendation systems

Sample DL Architectures (5)



DL Applications and Challenges (6)

- Criticisms
  - Manually designing deep models based on "trial and error"
  - Requiring a lot of domain knowledge
  - "A main criticism of deep learning concerns the lack of theory surrounding many of the methods. Most of the learning in deep architectures is just some form of gradient descent".
  - Deep learning methods are often looked at as a black box, with most confirmations done empirically, rather than theoretically.
  - Require very expensive computation – normal computers can't do – parallel computing, GPU, etc.
- Automated/Evolutionary Deep learning

Transfer Learning

- Supervised learning assuming a **similar** data distribution over Training set vs Test set
- When different, need to use transfer learning
- Transfer knowledge learned from one domain (source domain) to another domain (target domain)
  - Domain adaptation
  - Domain generalisation
  - Supervised vs unsupervised vs hybrid
- Multi-task learning
  - Multiple tasks to be solved simultaneously
  - Using a single representation
  - Knowledge sharing
- Representation learning

Further AI Courses in 2022

- AIML131: Introduction to AI (overview)
- COMP307: Introduction to AI (AI algorithms and techniques)
- COMP309: on how to use ML tools for problem solving
- AIML420: Artificial Intelligence (x COMP307, COMP420)
- AIML421: Machine Learning Tools & Techniques (X COMP309)
- AIML425: Neural Networks and Deep Learning
- AIML426: Evolutionary Computation and Learning
- AIML427: Big data
- AIML428: Text Mining and Natural Language Processing
- AIML429: Probabilistic Machine Learning
- AIML430: Applications and Implications of AI
- AIML431: Current Topics in Artificial Intelligence
- AIML440: Directed Individual Study
- AIML441: Directed Individual Study (30pt)
- AIML487: Research Project (45pt)
- AIML501: Research Essay in Artificial Intelligence
- AIML589: Research Project (45pt)
- AIML591: Master's thesis in AI
- AIML690/692/694: PhD thesis in AI

New AI Qualifications at VUW from 2021

- **Master of AI (MAI): 180pt**
  - **Entry**: completed a Bachelor's degree in computer science or a related subject with an average grade of at least B in the relevant final year courses
  - Requirements:

    Part 1 (120pt)
    - AIML430;
    - 45 further points from AIML 425–440;
    - 30 further points from AIML 420–489, ECEN 422, 430;
    - 30 further 400-level points from AIML, COMP, SWEN, NWEN, DATA, ECEN 422, 430, STAT 432, 452

    Part 2 (60pt)
    - AIML 501 and 589

New AI Qualifications at VUW from 2021

- **BSc (Hons) in Artificial Intelligence (AIML)**
  - **Entry**: COMP 307; 45 further points from (COMP 301-399, CYBR 301-399, NWEN 301-399, SWEN 301-399)
  - Requirements:
    - AIML 487 (45pt, research project)
    - 45 points from AIML 425–440;
    - 30 further 400-level points from AIML, COMP, SWEN, NWEN, DATA, ECEN 422, 430, STAT 432, 452
- **MSc in Artificial Intelligence (AIML)**
  - **Part 1 (120pt):**
    - AIML 430; 45 further points from AIML 425–440;
    - 30 further points from AIML 420–489, ECEN 422, 430;
    - 30 further 400-level points from AIML, COMP, SWEN, NWEN, DATA, ECEN 422, 430, STAT 432, 452
  - **Part 2**: AIML 591 (120pt, research thesis)

New AI Qualifications at VUW from 2021

- **PGDipSc in Artificial Intelligence (AIML)**
  - Requirements:
    - 45 points from AIML 425–440;
    - 30 further points from AIML 420–440, ECEN 422, 430;
    - 45 further 400-level points from AIML, COMP, DATA, SWEN, NWEN, ECEN 422, 430, STAT 432, 452 (including AIML 487)
- **PGCertSc in Artificial Intelligence (AIML)**
  - Requirements:
    - 60 points from AIML 400–489, ECEN 422, 430
- **PhD in Artificial Intelligence**
  - **3-4 years full time**

New AI Qualifications at VUW from 2022

- **BSc in Artificial Intelligence (AIML Major)**
  - **First AI major in NZ**
- **AIML Specialisations**
  - SWEN
  - EEEN
  - CYBE (working)

Research Scholarships at VUW

- Graduate Award: for BSc (COMP) Honours or MSc Part 1
  - $5000 covering tuition fees, Due on 1 Nov
- Summer Research Scholarships
  - $6000-8000, tax free, 400 hours, Due on 1 September or late August
  - Choose a research project
  - Can combine with DIS or BE Work experience
- Master by Research Scholarship (One year)
  - Covering local tuition fees
  - Stipend of $15,000 p.a.
- PhD Scholarships (three years)
  - Covering tuition fees
  - Stipend of $27,500 p.a.

Research Scholarships in AI (EC and ML)

- AI scholarships for BSc (AIML) Honours Research Project
  - Graduate Award + $1,000-2,000 (AI research grants)
- AI Summer Research Scholarships
  - $6,000-8,000 + $1,000, tax free, 400 hours, Due on 1 August
  - Choose an AI research project
  - Can combine with DIS or BE Work experience
- AI Master by Research Scholarship (One year)
  - Covering local tuition fees
  - Stipend of $15,000 + $2,000-3,000 p.a. (more than VUW ones)
- AI PhD Scholarships (three years)
  - Covering tuition fees
  - Stipend of $27,500 + 2,500-5,000 p.a. (more than VUW ones)
- Contact Meng (Mengjie Zhang) if you are interested

Potential AI Projects in Summer and Honours

- Machine Learning Techniques for Designing and Developing Mussel Mood Monitor (Meng, Bing with Cawthron)
- Predictability of Seafood Product Quality using Machine Learning Techniques (Bing, Meng with PFR)
- Multi-objective decision making for fin-fish production (Yi, Meng with PFR)
- Evolutionary Computer Vision and Image Processing (Meng/Bing/Harith)
- Evolutionary Scheduling and Routing (Yi/Fangfang/Meng)
- Evolutionary Transfer Learning (Bing/Bach/Yi/Fangfang/Meng)
- Automated deep learning for image classification and analysis (Bing, Meng)
- Modelling and symbolic regression with interpretable AI (Qi, Meng, Bing)
- Evolutionary Web Service Composition (Hui/Yi/Aaron)
- Deep learning for text mining and natural language processing (Sharon/Bing/Yi)
- Evolutionary clustering and unsupervised learning (Andrew/Meng/Bing)
- Resource allocation and planning (Hui/Yi/Aaron/Meng)
- Reinforcement learning and learning classifier systems (Aaron)
- Bayesian networks, Gaussian process, mathematical machine learning (Marcus/Bastiaan)
- Stream Data mining (Heitor)
- AI Ethics (Ali/Andrew)
- Maori/Pacifika AI (Bing/Yi/Meng)

Current Major AI Research Funding at VUW

- Data Science for Aquaculture, (AI/ML Advanced Research and Applications to Aquaculture). MBIE SSIF Fund on Data Science. 2020-2027. Grant: $13,000,000. (**Meng**; Ivy, Bing, Yi, Richard, Harith, Binh)
- Cyber-marine (AI/ML for Seafood with PFR). MBIE Research Program. 2020-2025. Grant: $16,800,000. (**Meng, Bing,** Ivy, Sharon, Qi, Aaron)
- Enabling unmanned aerial vehicles (drones) to use tools in complex dynamic environments. 2021-2026. MBIE Research Program. Grant: $9,837,002. (**Meng, Bing**)
- NS-TIP (AI/ML for Tree segmentation and species classification with Landcare). NZ-SQ Catalyst Fund. 2020-2023. Grant: $3,000,000. (**Meng, Bing**)
- Evolutionary Deep Learning for Image Classification. *Marsden Fund*. 2020-2023. Grant: $707,000. (**Bing**, Meng)
- GP for Symbolic Regression. *Marsden Fund*. 2020-2023. Grant: $707,000. (**Meng,** Qi/Bing)
- Interpretable GP for Symbolic Regression. Marsden Fund. 2021-2024. Grant: $360,000. (**Qi,** Bing, Meng)
- Precision Farming. *National Science Challenge SfTI Spearhead*. 2019-2022. Grant:$3,000,000 (Trench 2). VUW Grant: $512,300. **(Meng**, Bing)
- Large-scale Evolutionary Feature Selection for Classification. *Marsden Fund,* 2016-2022. Grant: $300,000. (**Bing**, Meng)
- Automatic Design of Heuristics for Dynamic Arc Routing Problem with Genetic Programming. *Marsden Fund,* 2016-2022. Grant: $300,000. (**Yi**, Meng)
- Genetic Programming for Dynamic Flexible Job Shop Scheduling. *Marsden Fund,* 2016-2022. Grant: $550,000. **(Meng)**
- Deep Learning Architecture with Context Adaptive Features for Image Parsing. *ARC Discovery Project*. Grant: $480,000 AUD. 2020-2022. (**Brijesh**, Meng)

VICTORIA UNIVERSITY OF WELLINGTON TE HERENGA WAKA

ECRG Evolutionary Computation Research Group

Artificial Intelligence, Machine Learning, and Data Science at VUW

Capabilities and Applications

Mengjie Zhang, FRSNZ, FEngNZ, FIEEE, IEEE DL

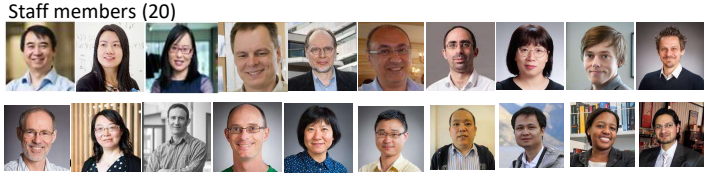
Artificial Intelligence/Data Science Research Cluster

Victoria University of Wellington, New Zealand

mengjie.zhang@ecs.vuw.ac.nz

Our People in DSAI/ML

- Staff members (20)

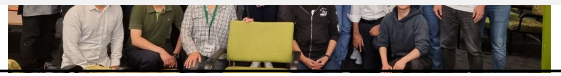


- Big A

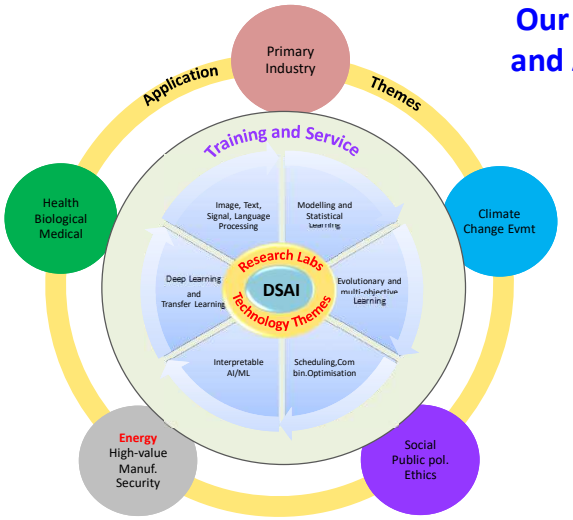
- PhD students (30)

Largest AI/ML team in NZ

Largest Research Group in evolutionary learning in Southern Hemisphere

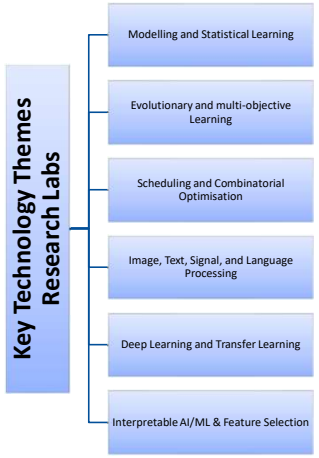


Our Capabilities and Applications



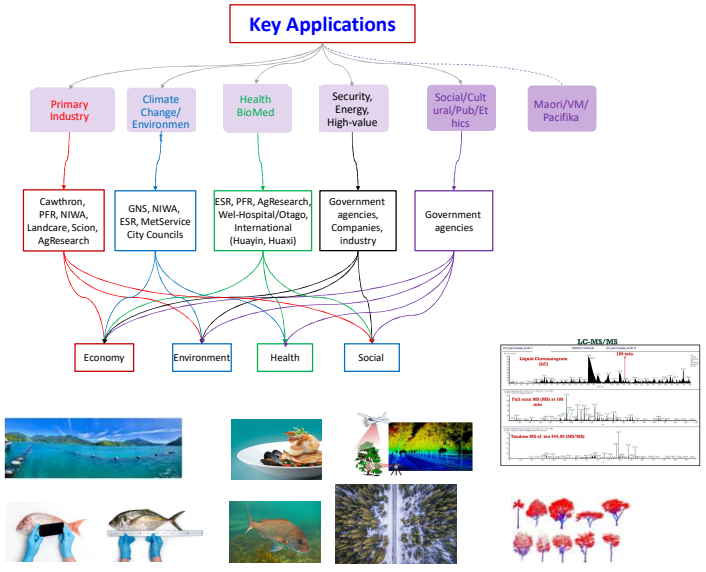
41

Major Technology Themes



- Have international leadership and reputation
- Potentially serve key applications
- Lead to substantial impact to economy, social, environmental and health outcomes

Key Applications



Summary

- Many interesting areas in AIML for students to choose
- Very popular in Industry
- Great opportunities for Research and Scholarships
- High salaries in job markets
- Many high-level jobs the Wellington and world require an MSc in AIML and PhD in AIML
- VUW provides
  - First AIML Postgraduate qualifications, and
  - First undergraduate major in AIML
- Contact mengjie.zhang@ecs.vuw.ac.nz