Advanced Artificial Intelligence Technologies and Applications

Course organiser: A/Prof. Shihua Zhou

Course presenter

Prof Nikola Kasabov

Visiting Professor at Dalian University

Life FIEEE, FRSNZ, FINNS, DVF RAE UK

Founding Director KEDRI

Professor, Auckland University of Technology, NZ

George Moore Chair/Professor, Ulster University, UK

Honorary Professor, University of Auckland NZ , Peking University China

Visiting Professor IICT/Bulgarian Academy of Sciences and Teesside University UK

Doctor Honoris Causa Obuda University Budapest

Director, Knowledge Engineering Consulting Ltd (https://www.knowledgeengineering.ai)

Assistants

A/Prof. Wei Qi Yan

Director of the CeRV Center, AUT

Weiqi.yan@aut.ac.nz

https://academics.aut.ac.nz/weiqi.yan

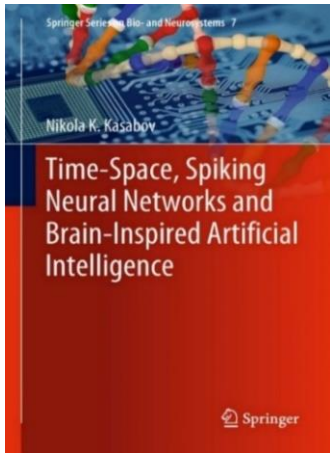
Ms Iman AbouHassan

iabouhassan@tu-sofia.bg

abouhassan.iman@gmail.com

Advanced Artificial Intelligence Technologies and Applications

1. AI and the evolution of its principles. Evolving processes in Time and Space (Ch1, 3-19)
2. From Data and Information to Knowledge. Fuzzy logic. (Ch1,19-33 + extra reading)
3. Artificial neural networks - fundamentals. (Ch2, 39-48). Computational modelling with NN. Tut1: NeuCom.
4. Deep neural networks (Ch.2, 48-50 + extra reading).
5. Evolving connectionist systems (ECOS) (Ch2, 52-78). Tutorial 2: ECOS in NeuCom.
6. Deep learning and deep knowledge representation in the human brain (Ch3)
7. Spiking neural networks (Ch4). Evolving spiking neural networks (Ch5)
8. Brain-inspired SNN. NeuCube. (Ch.6). Tutorial 3: NeuCube software (IA)
9. From von Neuman Machines to Neuromorphic Platforms (Ch20 , 22)
10. Other neurocomputers: Transformers.
11. Evolutionary and quantum inspired computation (Ch.7)
12. AI applications for brain data: EEG, fMRI (Ch.8-11)
13. Brain-computer interfaces (BCI) (Ch.14)
14. AI applications for audio-visual information (Ch.12,13). AI for language modelling.
15. AI in bioinformatics and neuroinformatics (Ch15,16, 17,18)
16. AI applications for multisensory environmental data (Ch19).
17. AI in finance and economics (Ch19)



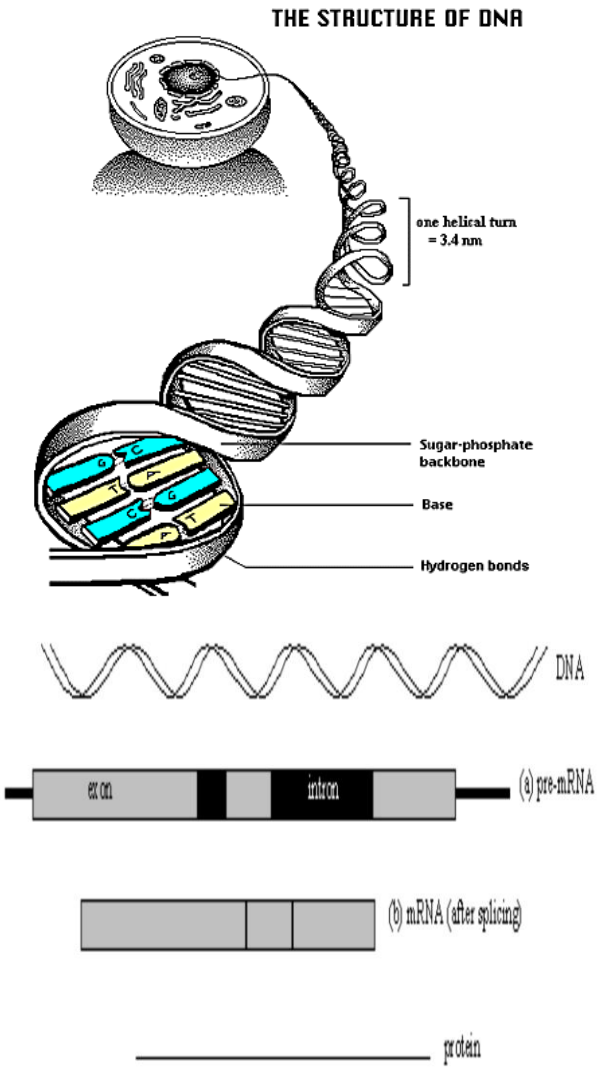
Course book: N.Kasabov, Time-Space, Spiking Neural Networks and Brain-Inspired Artificial Intelligence Springer, 2019, https://www.springer.com/gp/book/9783662577134

Additional materials: https://www.knowledgeengineering.ai/china

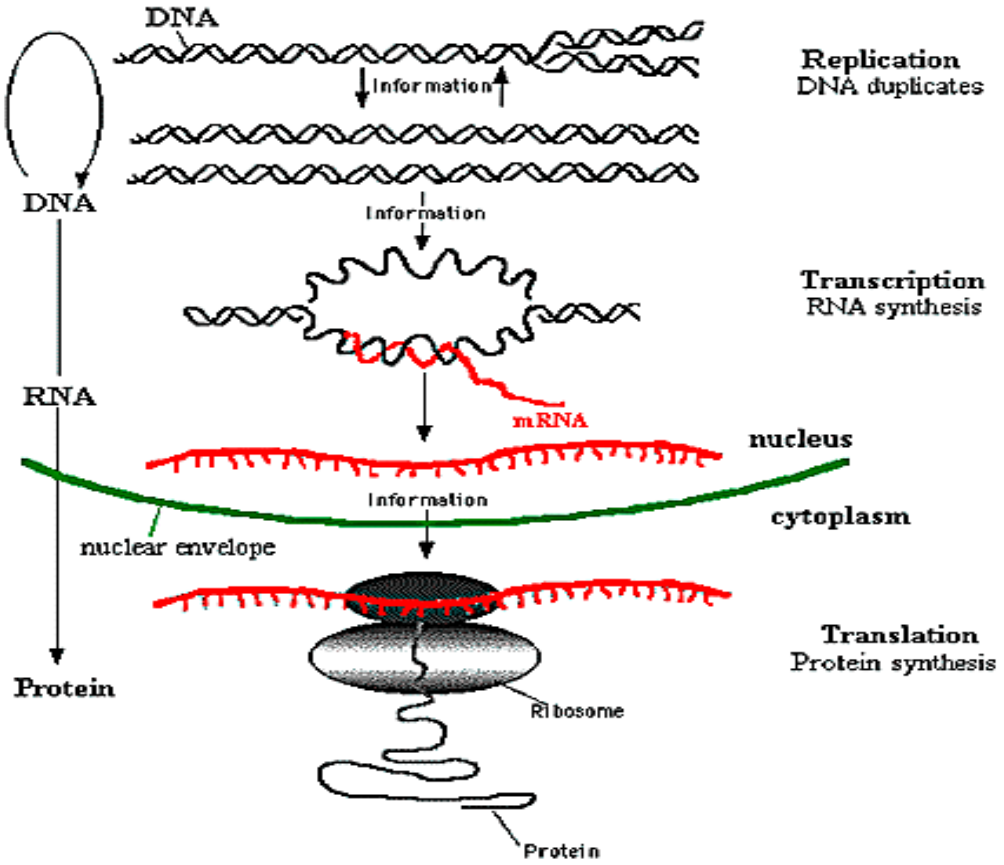
ZOOM link for all lectures: https://us05web.zoom.us/j/4658730662?pwd=eFN0eHRCN3o4K0FaZ0lqQmN1UUgydz09

DNA, Genes and Proteins (Ch.2)

- Each cell of an individual contains the whole DNA (the genome) of the individual.
- About 36,000 genes in the human genome, each of them comprising of 50 to a mln base pairs – A,T,C or G – basic molecules)
- The Main Dogma: DNA->RNA->proteins
  - Transcription: DNA (about 5%) -> mRNA
    - DNA -> pre-RNA -> splicing -> mRNA (only the exons)



Chapter 15. Computational modelling and pattern recognition in Bioinformatics

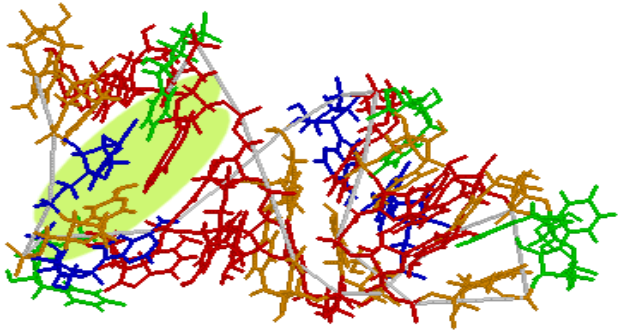


The Central Dogma of Molecular Biology

mRNA is the only type of RNA that is translated into protein

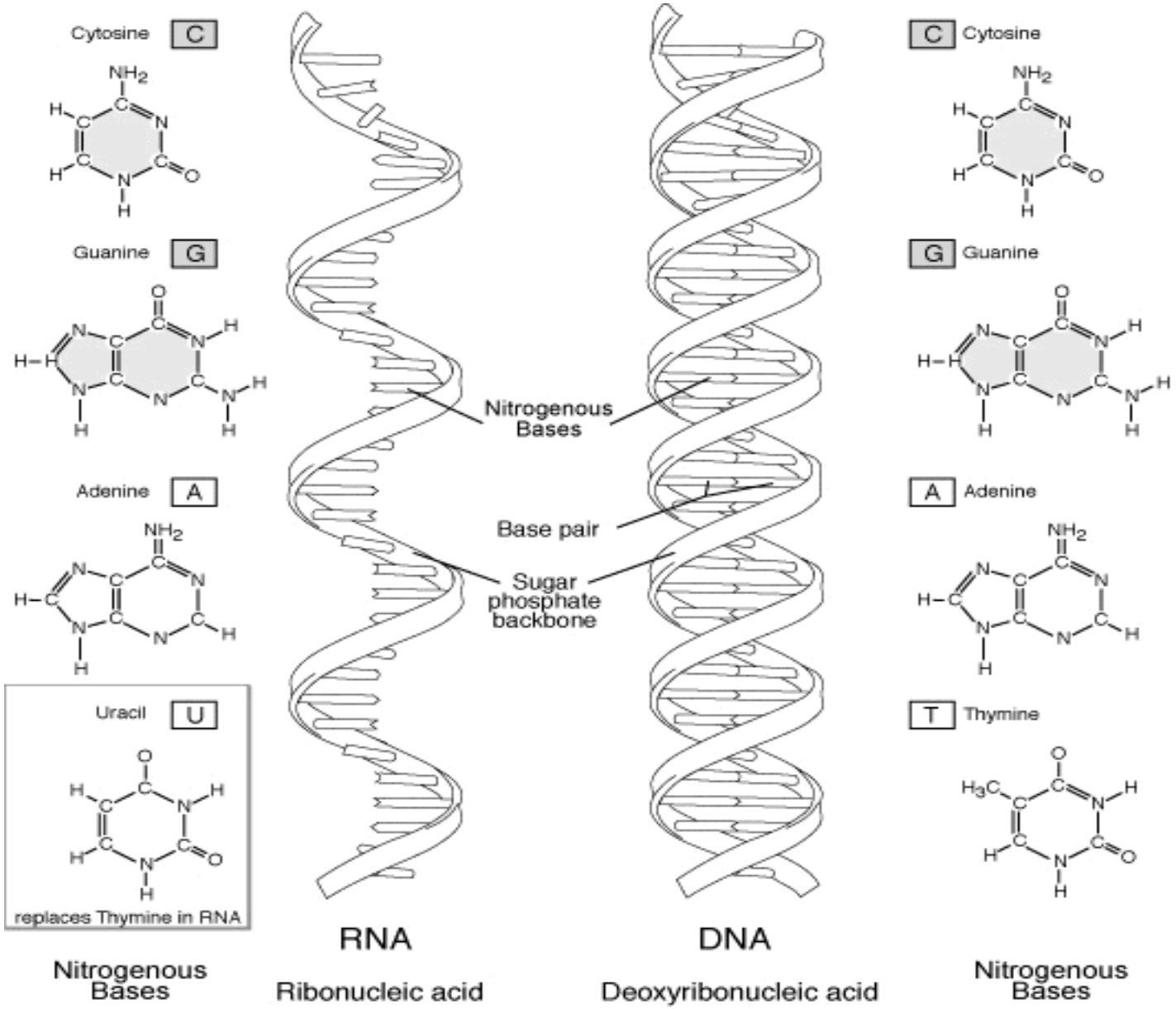
DNA → (Transcription) → RNA

1)tRNA (9%)
2)rRNA (85%)
3)snRNA (1%)
4)mRNA (5%)

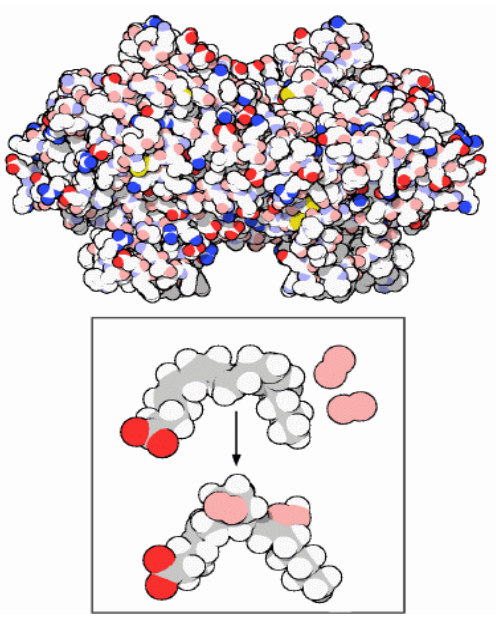


→ (Translation) → Protein

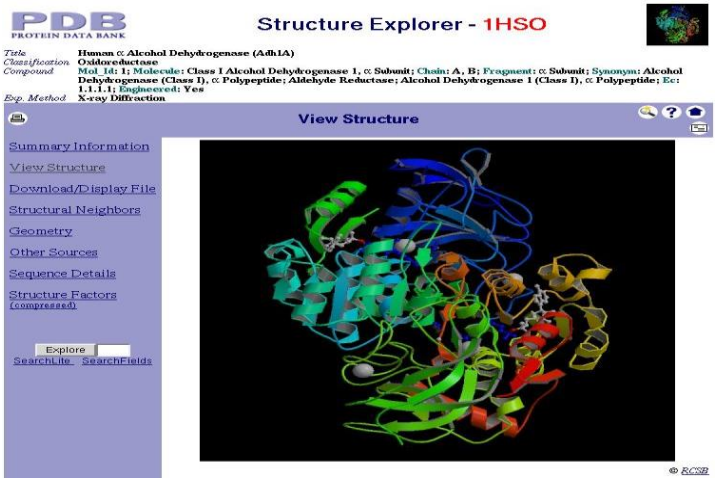
Chemistry of DNA and RNA



Proteins and protein structures

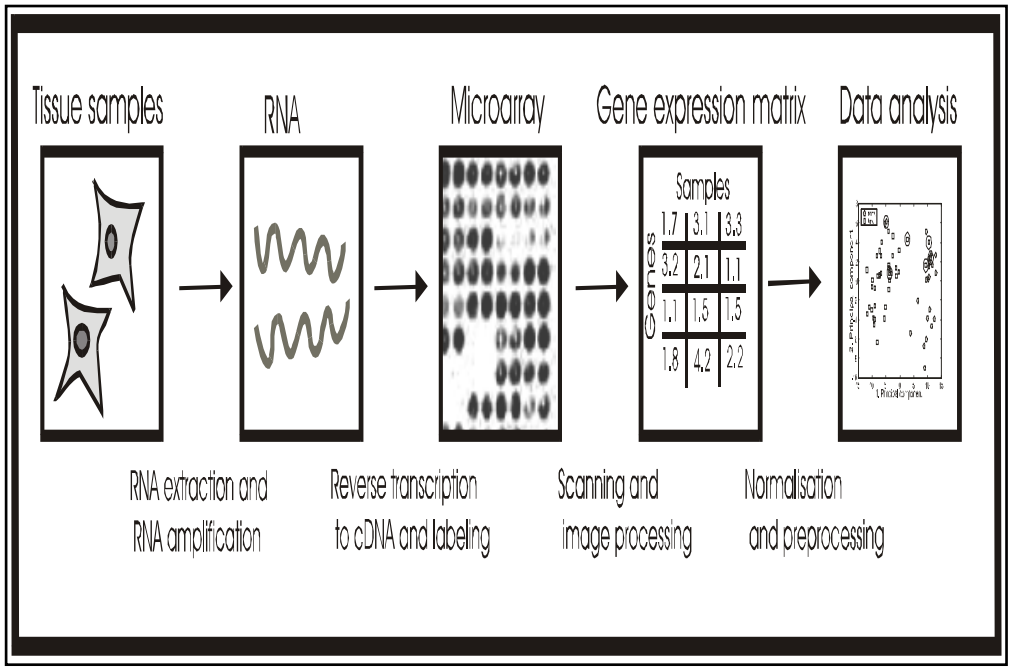


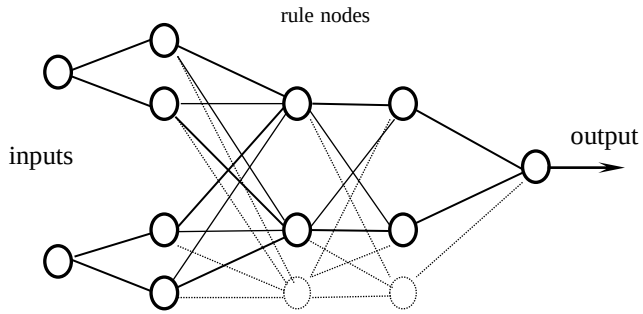
- The mRNA is translated into proteins
- A protein is a sequence of amino-acids, each of them defined by a group of 3 nucleotides (codons)
- 20 amino acids all together (A,C-H,I,K-N,P-T,V,W,Y)
- Initiation and stop codons
- Proteins have complex structures:
  - Primary (linear),
  - Secondary (3D, defining functionality)
  - Tertiary (energy minimisation packs),
  - Quaternary (interaction between molecules)
- The Protein Data Bank – www.rcsb.org - 100,000 hits a day on average



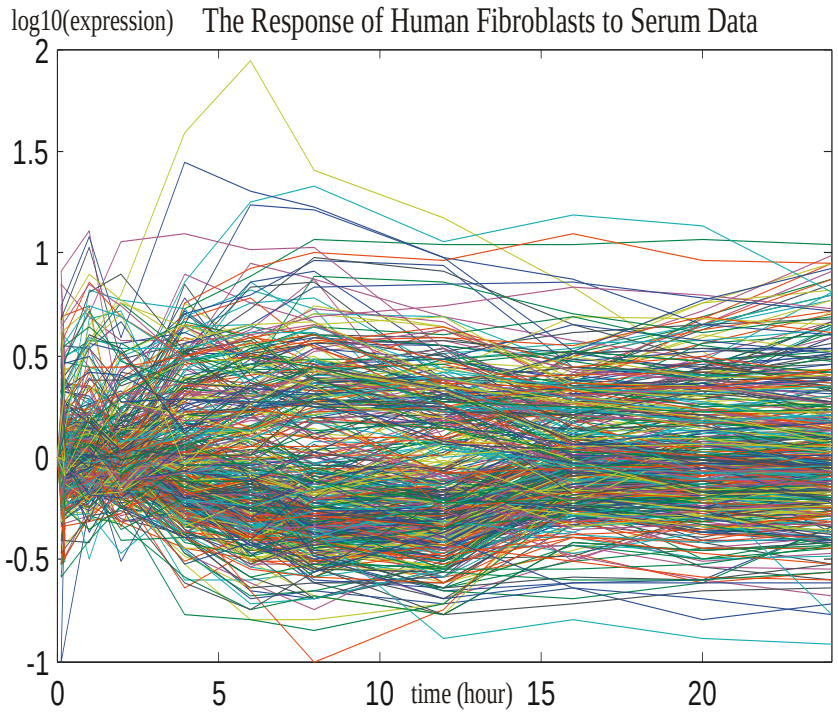
Gene expression data and their modelling

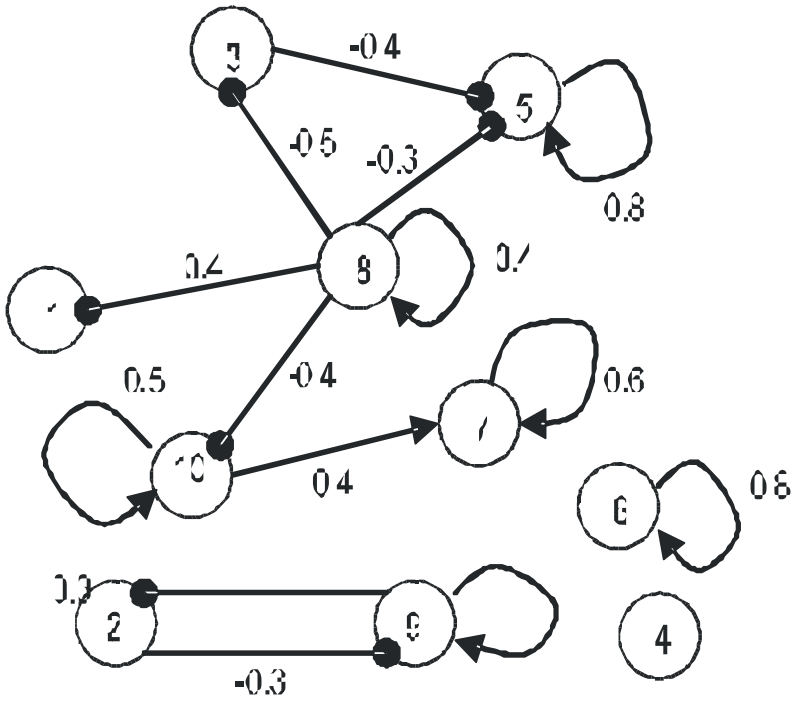
- Gene expression data analysis
- Goal: identify a gene or a group of genes associated with the state of the cell (tissue), e.g. cancer.
- Large number of genes (appr. 30,000) expressed in a microarray (in vitro) from a single tissue.
- It is difficult to find consistent patterns of gene expression for a class of tissue
- After all, a microarray data is just of few microseconds snapshot of what is happening in the cell
- Genes interact – how do we find out about that?
- Growing number of examples and complexity.



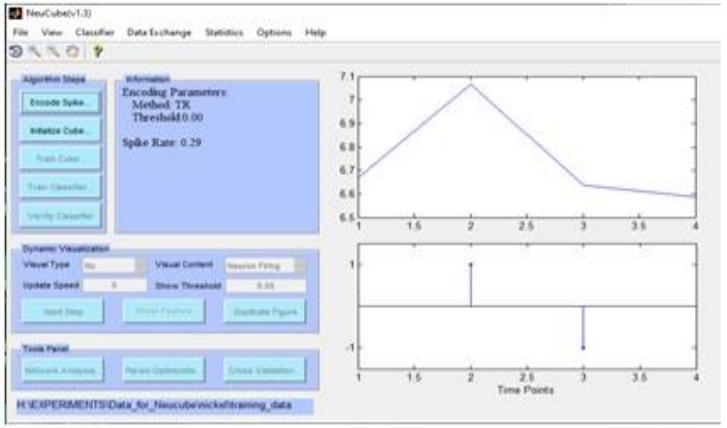


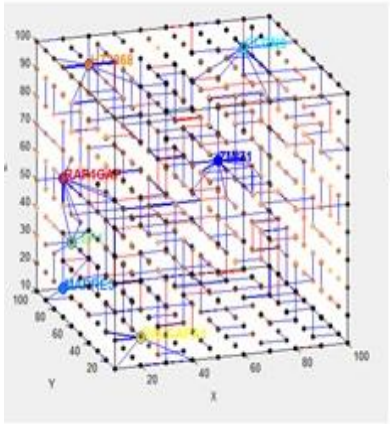
Gene expression time series data and gene regulatory networks (GRN) modelling

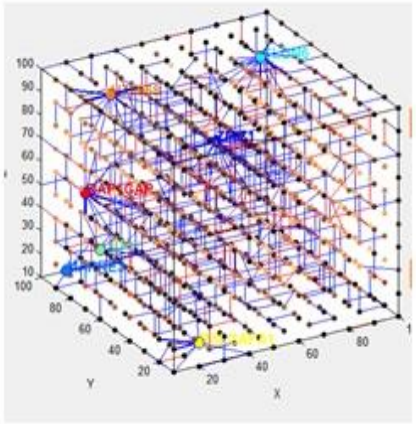


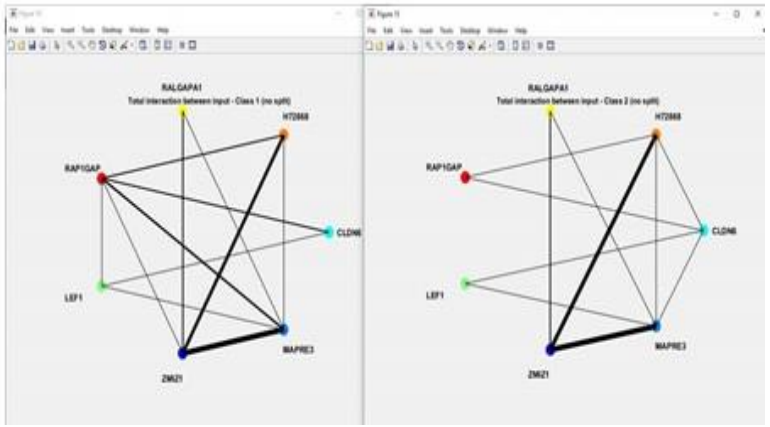


Example of modelling gene expression time series in NeuCube







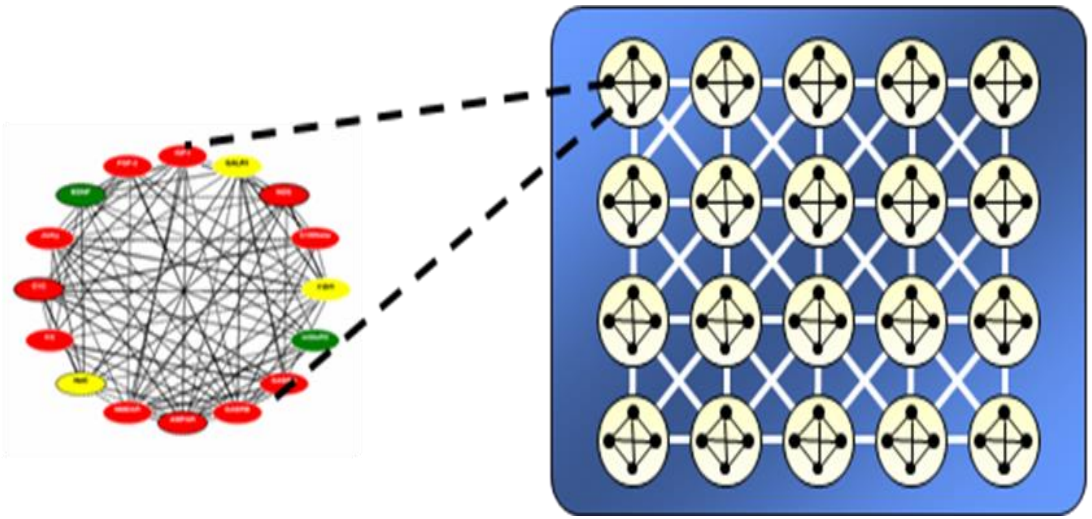


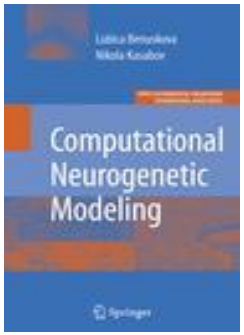
Chapter 16. Computational Neuro-Genetic Modelling (CNGM)

- Benuskova and Kasabov (2007)

SNN that incorporate a gene regulatory network (GRN) as a dynamic parameter systems to capture dynamic interaction of genes (parameters) related to neuronal activities of the SNN.

- Functions of neurons and neural networks are influenced by internal networks of interacting genes and proteins forming an abstract GRN model.
- The GRN and the SNN function at different time scales.





A neurogenetic model of a spiking neuron

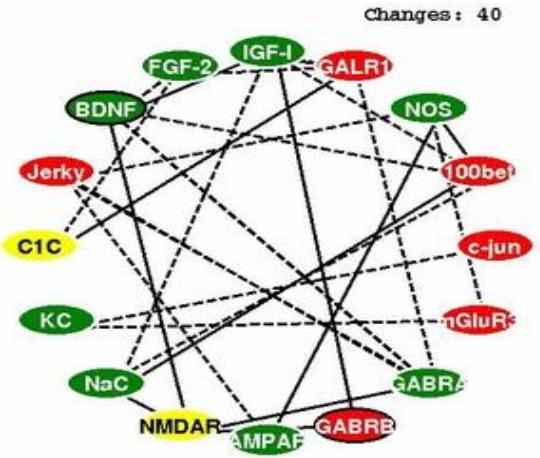
(Kasabov, Benuskova, Wysoski, 2005)

- Four types of synapses: fast excitation; slow_excitation; fast_inhibition; slow_inhibition

- A Gene Regulatory Network (GRN) as a dynamical parameter system of the neuron

Table. Neuronal Parameters and Related Proteins

Neuronal parameter Amplitude and time constants of	Protein
Fast excitation PSP	AMPAR
Slow excitation PSP	NMDAR
Fast inhibition PSP	GABRA
Slow inhibition PSP	GABRB
Firing threshold	SCN, KCN, CLC
Late excitatory PSP through GABRA	PV



$$PSP_{ij}^{type}(t - t_j - \Delta_{ij}^{ax}) = A^{type}\left(\exp\left(-\frac{t - t_j - \Delta_{ij}^{ax}}{\tau_{decay}^{type}}\right) - \exp\left(-\frac{t - t_j - \Delta_{ij}^{ax}}{\tau_{rise}^{type}}\right)\right)$$

type = fast excitation; slow_excitation; fast_inhibition; slow_inhibition

nkasabov@aut.ac.nz

An example of a derived GRN through CNGM: A case study on epilepsy (with A. Villa et al, 2006))

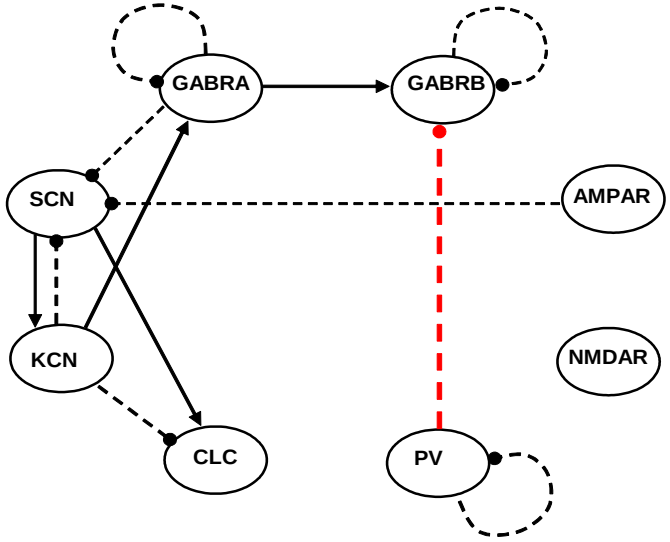
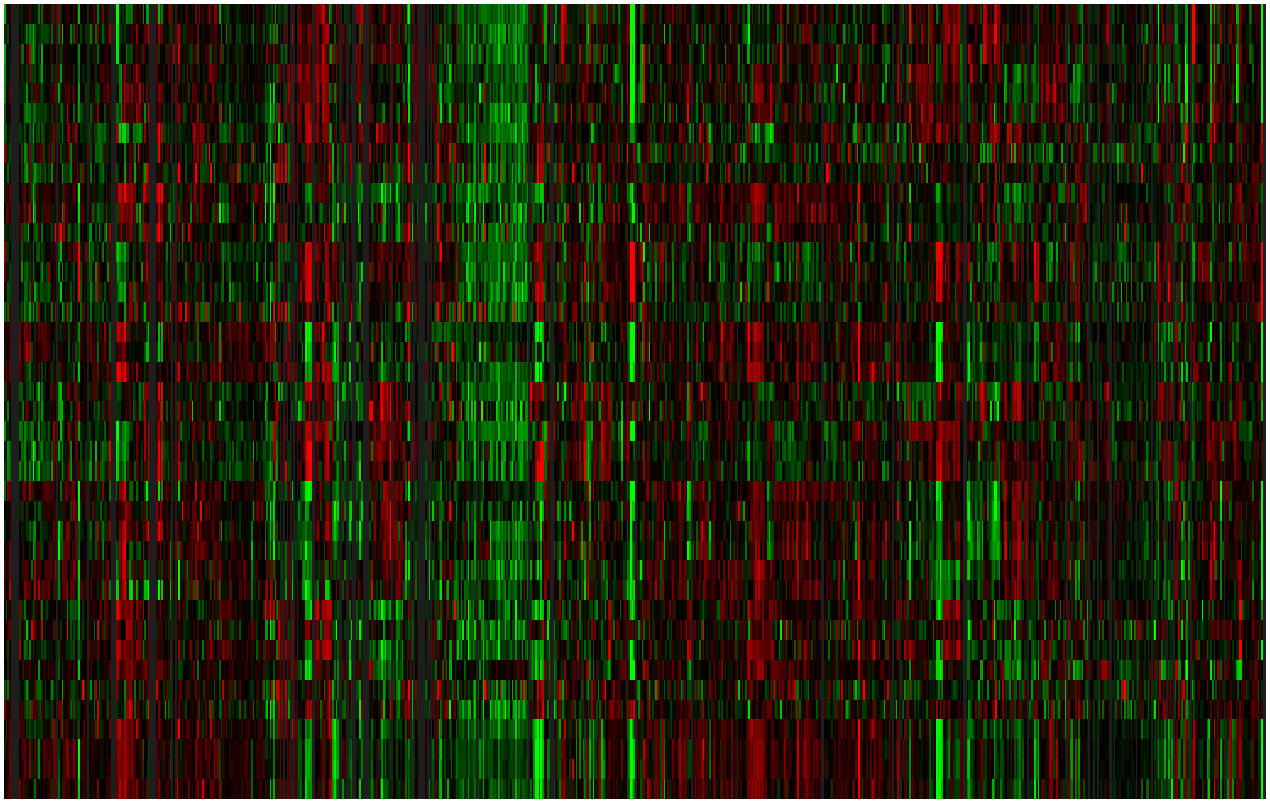


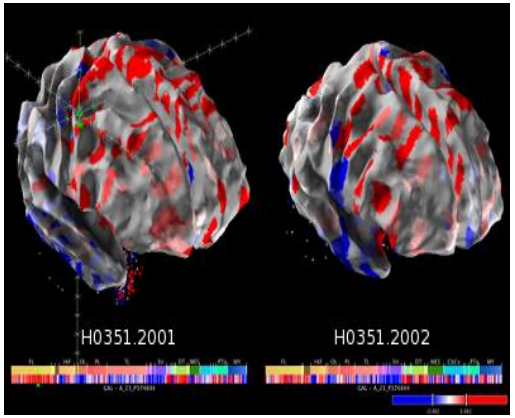
Table. Neuronal Parameters and Related Proteins

Neuronal parameter Amplitude and time constants of	Protein
Fast excitation PSP	AMPAR
Slow excitation PSP	NMDAR
Fast inhibition PSP	GABRA
Slow inhibition PSP	GABRB
Firing threshold	SCN, KCN, CLC
Late excitatory PSP through GABRA	PV

- The strongest interactions between genes in the evolved abstract GRN (left picture)
- The GRN can be used to predict gene knock-out consequences
- Predicted consequence of PV gene knock-out upon expression of GABRB – stronger slow inhibition
- Predicted consequence of PV gene knock-out upon local field potential – shift to lower frequencies of oscillations
- Potential for modeling of epilepsy and other genetic diseases manifested by the change of EEG/LFP

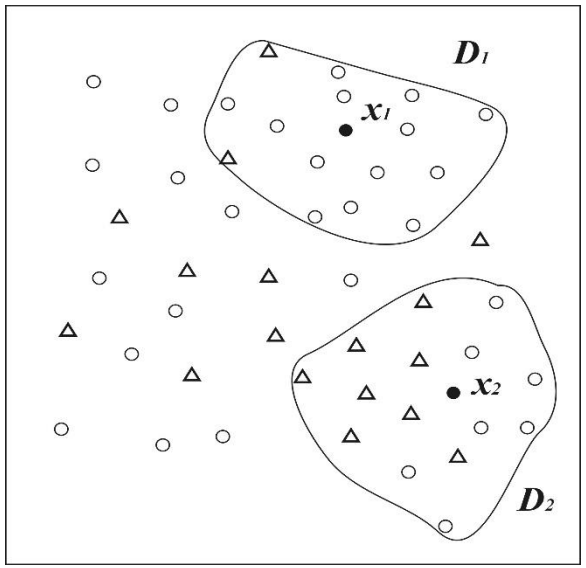
Neurogenetic STBD: The Allen Brain Institute Map (http://www.brain-map.org)





From the Brain Explorer: The Expression level of the genes (on the y-axis): ABAT A_23_P152505, ABAT A_24_P330684, ABAT CUST_52_PI416408490, ALDH5A1 A_24_P115007, ALDH5A1 A_24_P923353, ALDH5A1 A_24_P3761, AR A_23_P113111, AR CUST_16755_PI416261804, AR CUST_85_PI416408490, ARC A_23_P365738, ARC CUST_11672_PI416261804, ARC CUST_86_PI416408490, ARHGEF10 A_23_P216282, ARHGEF10 A_24_P283535, ARHGEF10 CUST_) at different slices of the brain (on the x-axis) (from www.brain-map.org) (http://www.alleninstitute.org)

Chapter 17. Computational framework for personalised modelling. Applications in Bioinformatics



- A transductive model is created on a sub-set of neighbouring data to each input vector. A new data vector is situated at the centre of such a sub-set (here illustrated with two of them – x_1 and x_2), and is surrounded by a fixed number of nearest data samples selected from the training data *D* and generated from an existing model *M (Vapnjak)*

- The principle of "What is good for my neigbours is good for me"

- Problems:

  - **Which variables, weighted or not weighted ?**
  - **How many neighbours?**
  - **What distance measure?**
  - **Which model?**

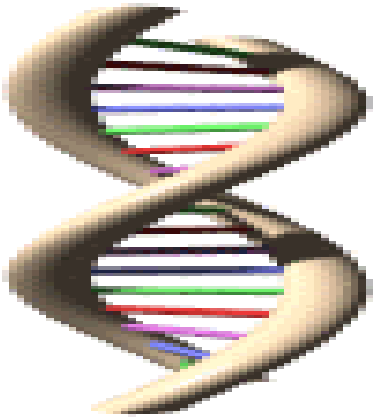
Parameter and feature optimisation:

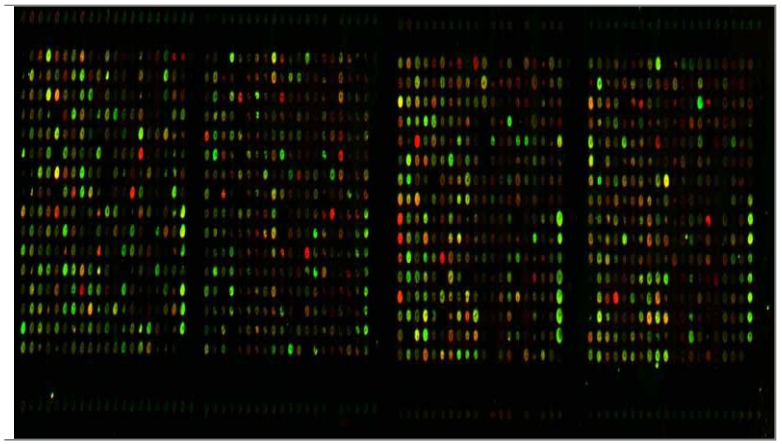
- GA (Proc. IJCNN05, IEEE Press, Montreal, 2005)

- LMS (e.g. TWNFI)

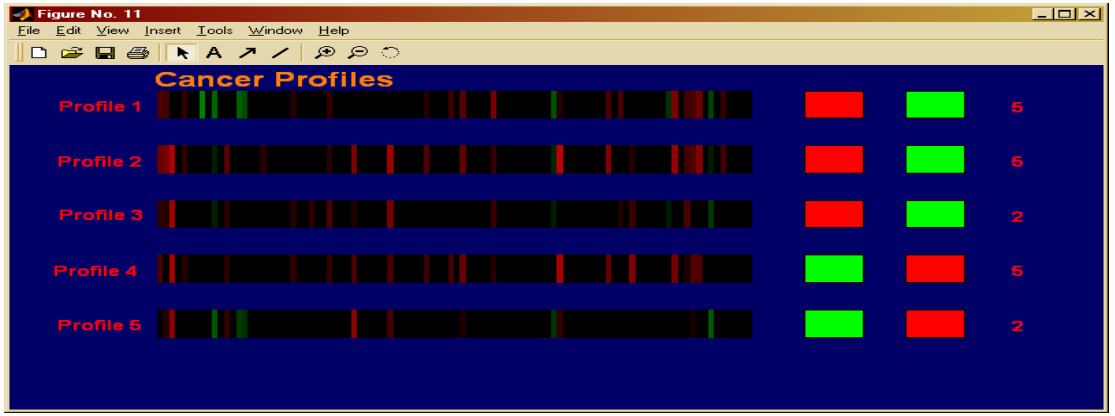
Personalised diagnosis and prognosis of cancer based on gene expression and clinical data

- DNA analysis - large data bases; data always being added and modified; different sources of information





- Cancer Ontology-Based DSS
- Markers and drug discoveries
- PEBL: www.peblnz.com

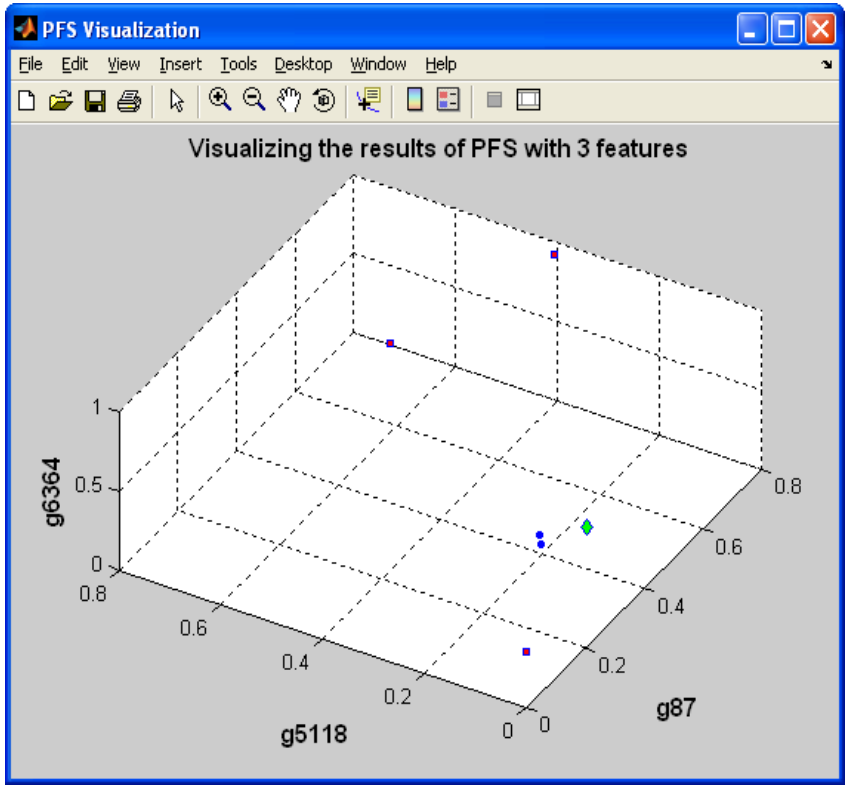


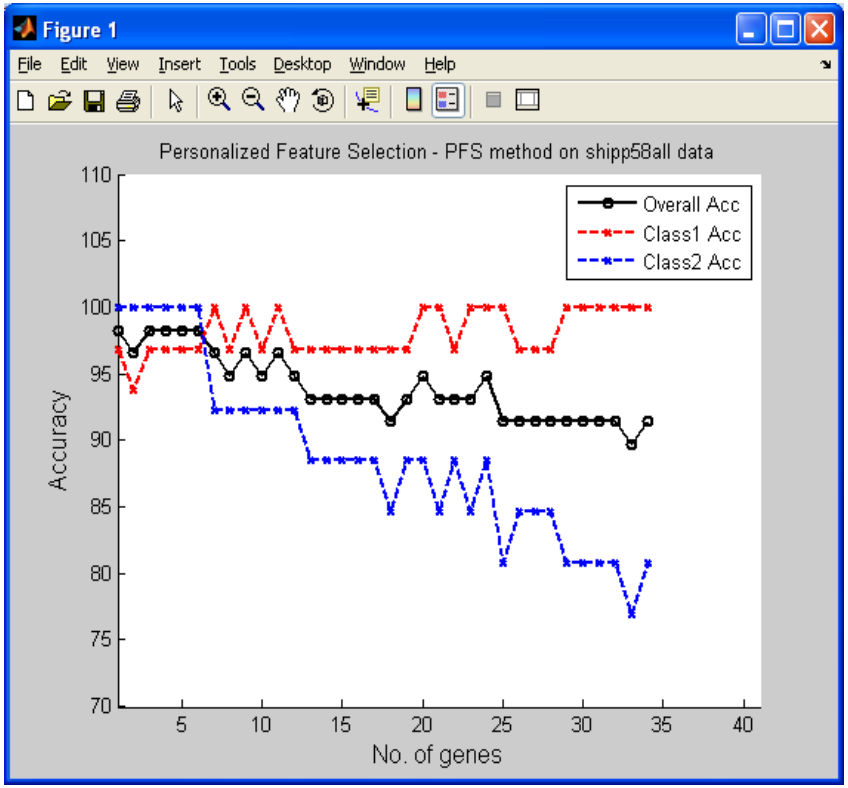
- Kasabov, N., Global, local and pesonalised m*odelling and profile discovery in Bioinformatics,* Pattern Recognition Letters, Jan. 2007

nkasabov@aut.ac.nz

Personalised modelling of gene expression data for cancer outcome prediction

(experiments on M.Shipp's data of DLBCL)





nkasabov@aut.ac.nz

Global, local and personalised modelling – a comparative study on the GFR example

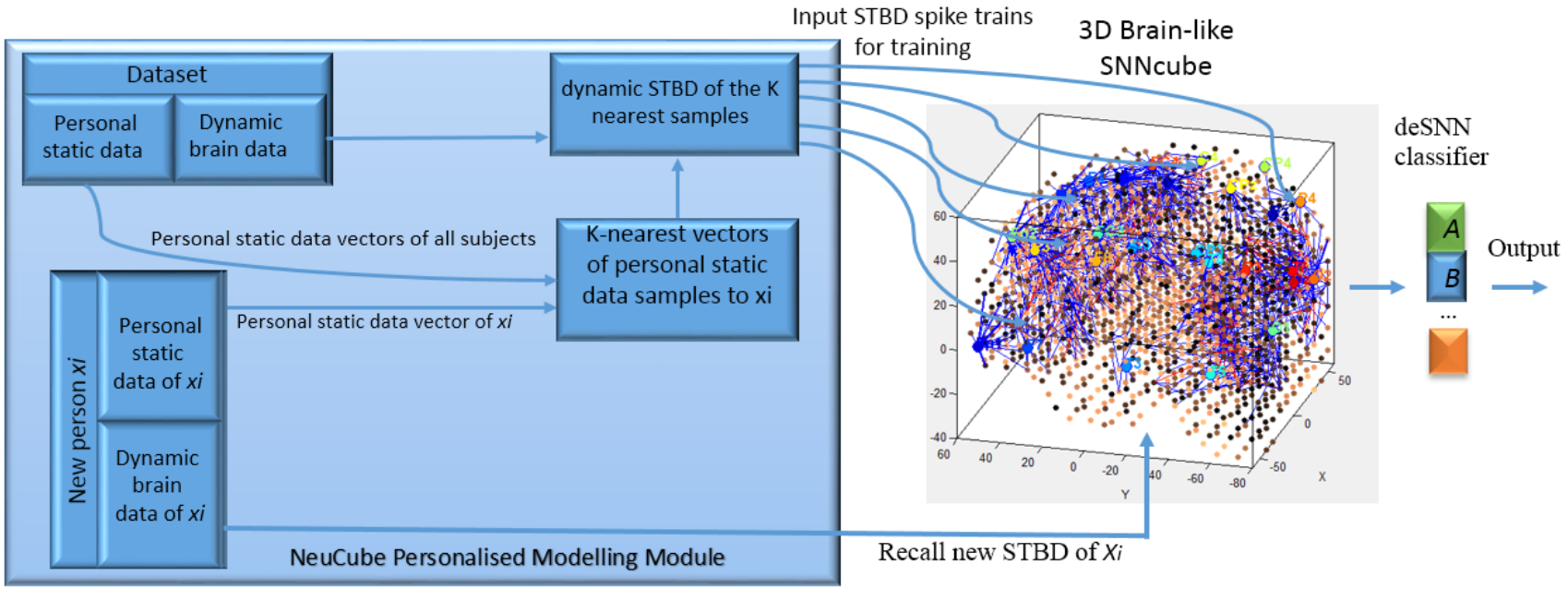
Using different models reveals different knowledge

Model	Neurons or rules	Test RMSE	Test MAE	Average weights of input variables					
				Age w1	Sex w2	Scr w3	Surea w4	Race w5	Salb w6
MDRD *(global)*	—	7.74	5.88	1	1	1	1	1	1
MLP (*global)*	12	8.38	5.71	1	1	1	1	1	1
ANFIS (*global)*	36	7.40	5.43	1	1	1	1	1	1
DENFIS (*local*)	27	7.22	5.21	1	1	1	1	1	1
TNFI *(personalised)*	**6.8 (average)**	**7.28**	**5.26**	**1**	**1**	**1**	**1**	**1**	**1**
TWNFI (*personalised)*	**6.8 (average)**	**7.08**	**5.12**	**0.87**	**0.70**	**1**	**0.93**	**0.40**	**0.52**

Chapter 18. Personalised modelling for integrated static and dynamic data. Applications in neuroinformatics.

Doborjeh, M., and Kasabov, N., IEEE WCCI/IJCNN, 2016 (Response to treatment of drug addicts using clinical and EEG data)

M. Doborjeh, N. Kasabov, Z. Doborjeh, R. Enayatollahi, E. Tu, A. H. Gandomi, Personalised modelling with spiking neural networks integrating temporal and static information, Neural Networks, 119 (2019),162-177.

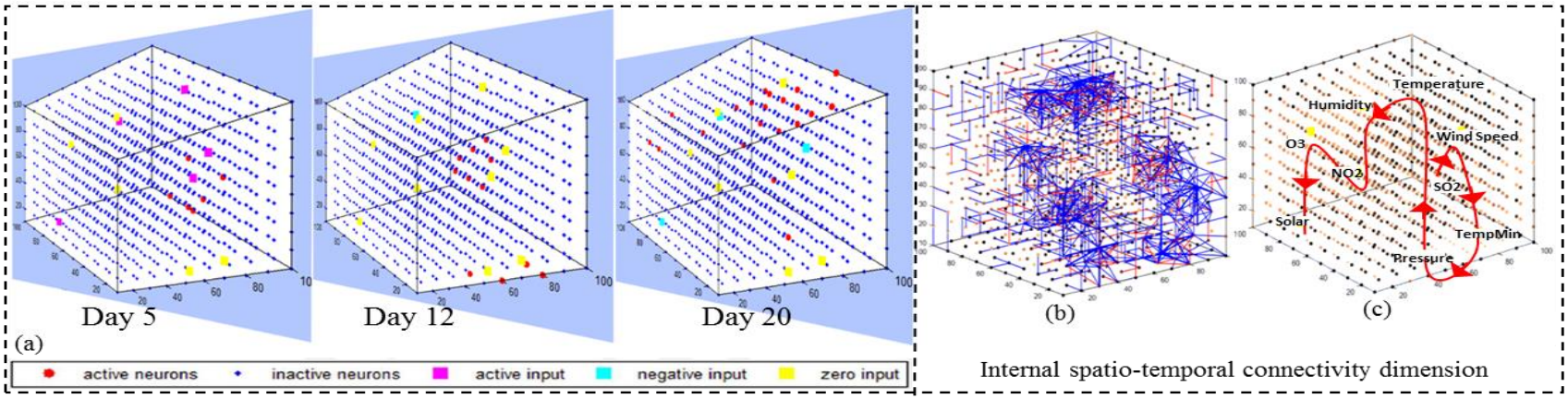


Methods	NeuCube-Personalised modelling	NeuCube- Global modelling
Classification accuracy of class M versus class OP in %	Averaged over 47 trained PSNN models: 93.61	One trained SNN model using all subjects and tested via leave-one-out method: 79.00

Personalised stroke prediction

Kasabov, N., Feigin, V., Hou, Z. -G., Chen, Y., Liang, L., Krishnamurthi, R., Parmar, P. (2014). Evolving spiking neural networks for personalised modelling, classification and prediction of spatio-temporal patterns with a case study on stroke. Neurocomputing, 134, 269-279. doi:10.1016/j.neucom.2013.09.049

Three snapshots of a NeuCube model during training on temporal climate and air pollution data of 9 variables, measured on each of 20 days before a stroke event happened to patients from a selected group (the left 3 figures). The evolved connectivity in the 3D SNN model after training – spatio-temporal structural patterns of connections are learned in the 3D dimensionality of the model. A dynamic functional pattern learned in the functional space of climate variable changes (the right most figure).



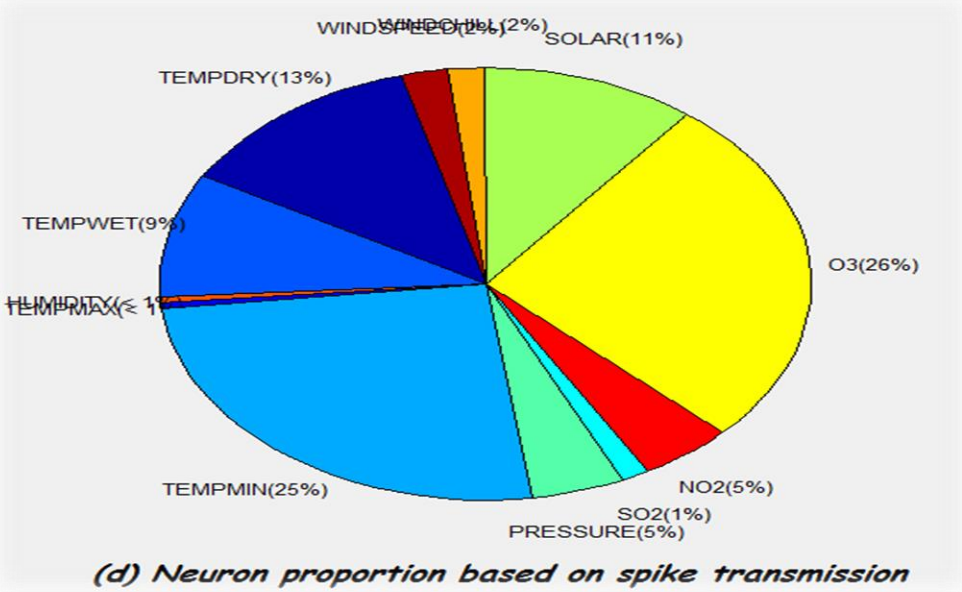
A spatio-temporal rule extracted from a trained SNNcube on climate data relate to a high risk of stroke for a group of individuals

IF SO2 changes around time T1) AND (Wind Speed changes around time T2)
AND (TempMin changes around time T3) AND (Pressure changes around time T4)
AND (AvTemp changes around time T5) AND (Humidity changes around time T6)
AND (NO2 changes around time T7) AND (O3 changes around time T8) AND (Solar eruption around T9)
THEN (High risk of stroke for the individual X and the group she/he belongs to)

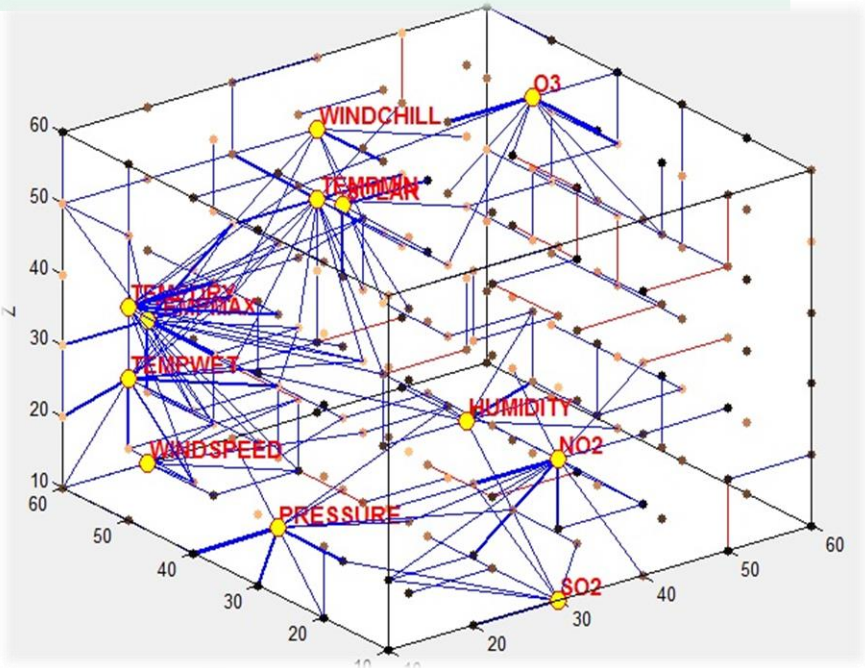
Personalised prediction of risk for stroke days ahead

(N.Kasabov, M. Othman, V.Feigin, R.Krishnamurti, Z Hou et al - Neurocomputing 2014)

Methods	*SVM*	*MLP*	*KNN*	*WKNN*	*NeuCubeST*
1 day earlier (%)	55 *(70,40)*	30 *(50,10)*	40 *(50,30)*	50 *(70,30)*	**95** ***(90,100)***
6 days earlier (%)	50 *(70,30)*	25 *(20,30)*	40 *(60,20)*	40 *(60,20)*	**70** ***(70,70)***
11 days earlier (%)	50 *(50,50)*	25 *(30, 20)*	45 *(60,30)*	45 *(60,30)*	**70** ***(70,70)***



(d) Neuron proportion based on spike transmission



- SNN achieve better accuracy
- SNN predict stroke much earlier than other methods
- New information found about the predictive relationship of variables

nkasabov@aut.ac.nz

Questions

1. What is the central dogma in biology?

2. How can NN be used for gene expression data classification?

3. How can NeuCube be used for modelling gene expression time series?

4. What is personalised modelling?

5. Give examples of personalised modelling in neuroinformatics.