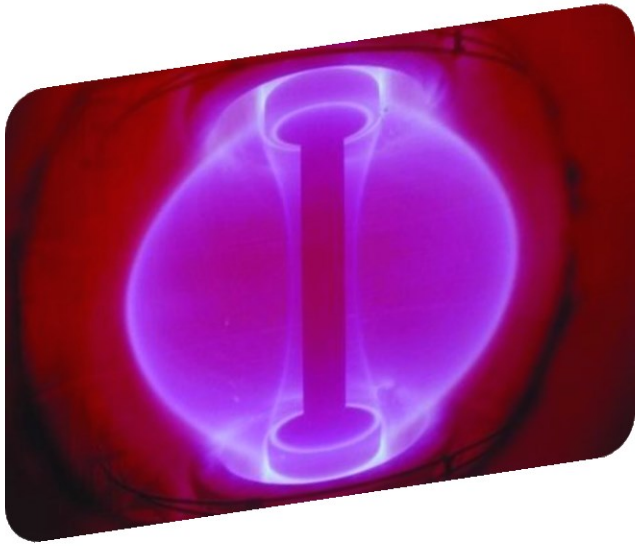


Machine Learning in Plasma Physics

Yavar Taheri Yeganeh

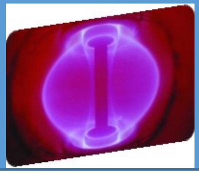
Department of Physics

Shahid Beheshti University (SBU)

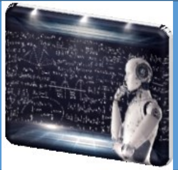
Plasma Physics Seminar @ SBU Physics

October 15, 2019

Please Note: References are not listed in the slides



Introduction & Content



- Who am I ?
- My Current Research
- Today Topic > Exciting & Stimulating

Quick detailed Introduction to the ML Approach

- Previous Talk Review
- AI Review
- Applications in Experimental Research
- Applications in Theoretical Research
- Conclusion > Discussion

YAVAR TAHERI YEGANEH

y.taheriyeganeh@mail.sbu.ac.ir
yavaryeganeh@gmail.com
(+98)9380493329
LinkedIn

Department of Physics
Shahid Beheshti Univerity
Evin, Tehran 19839, Iran

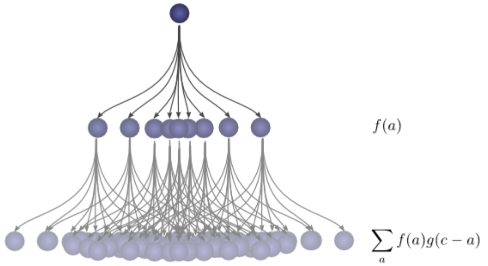
Education

MSc in Physics, Department of Physics, Shahid Beheshti University, Iran
Thesis title: 'Study on absorption of electromagnetic waves near UH resonance in high-beta plasma'
Supervisor: Prof. Mohammad Ghorbanalilu

BSc in Robotics Engineering (Machine Intelligence & Robotics),
Department of Electrical and Robotics Engineering, Shahrood University of Technology, Iran
Thesis title: 'Design, study and construction of magnetic robot'
Supervisor: Dr. Masoud Mahdizadeh

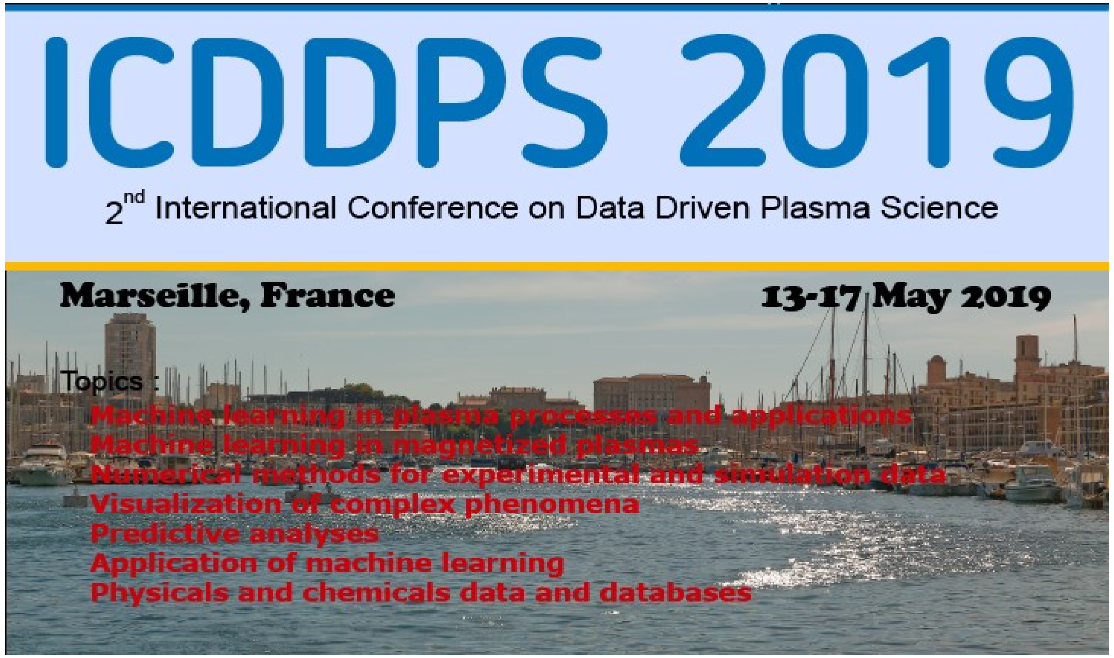
Areas of Interest

Plasma Physics • Artificial Intelligence • Machine Learning and Data Science • Computational Physics
Numerical Simulation • Control Theory • Plasma Diagnostics • High Performance Computing
Many-Body Physics • Experimental Particle Physics • Quantum Electrodynamics



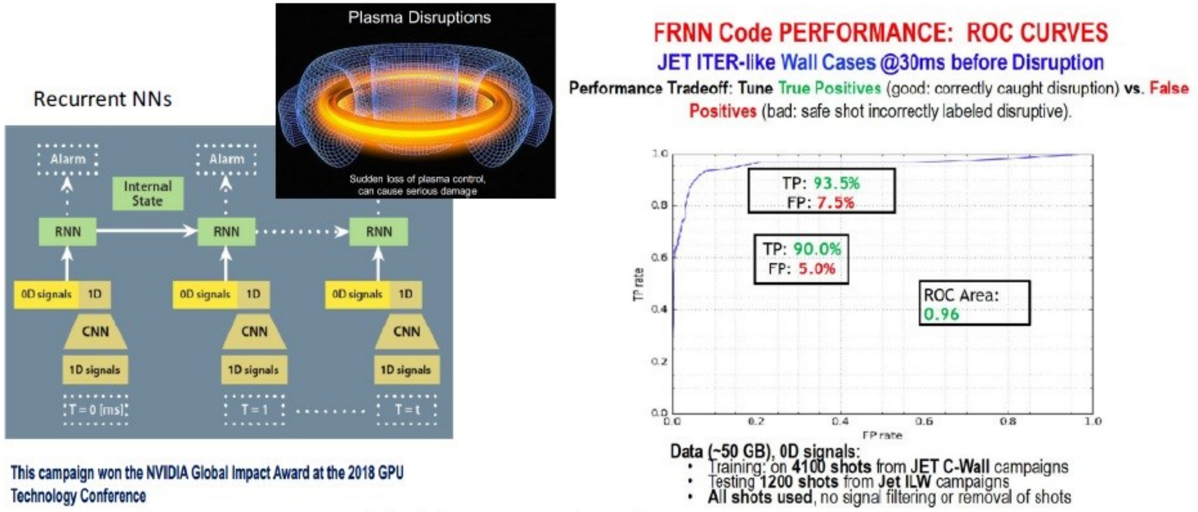
- AI Revolution and Plasma
- Brief Introduction on AI and DS
- AI and DS in Research

-> Computational Advancements in Plasma Physics



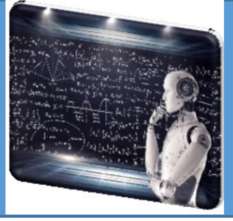
Similarity to financial data analysis, earthquake prediction etc.

DL for real-time disruption prediction



W. Tang et al. (talk by D. Keyes @ WS1)

AI Review



CLASSIFICATION → SEGMENTATION

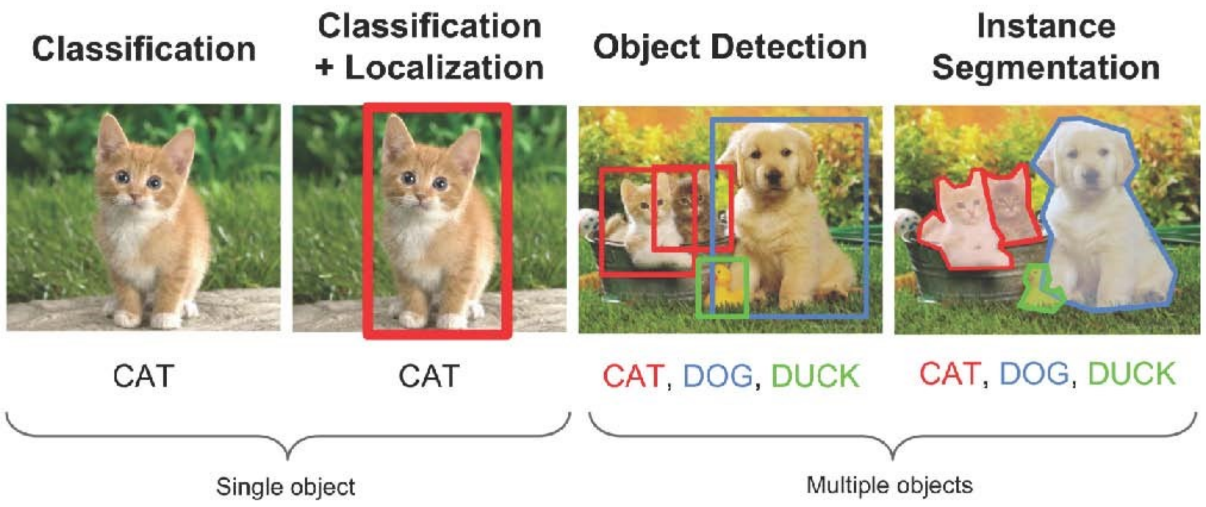
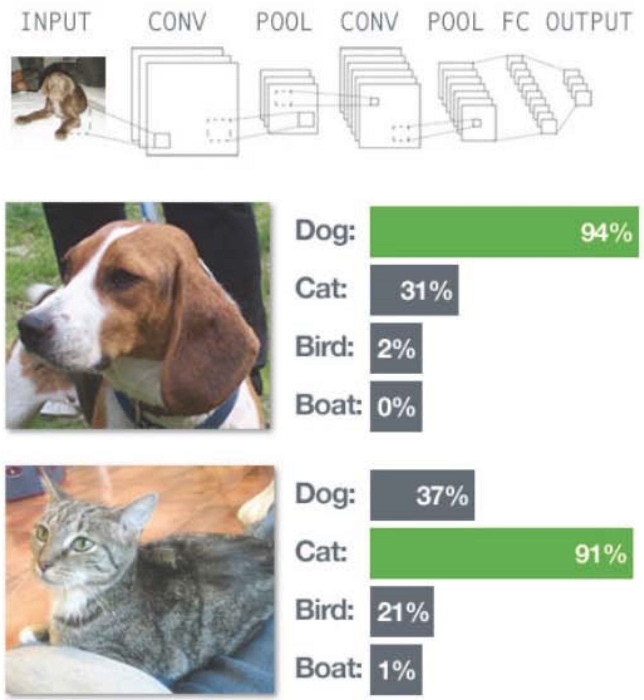
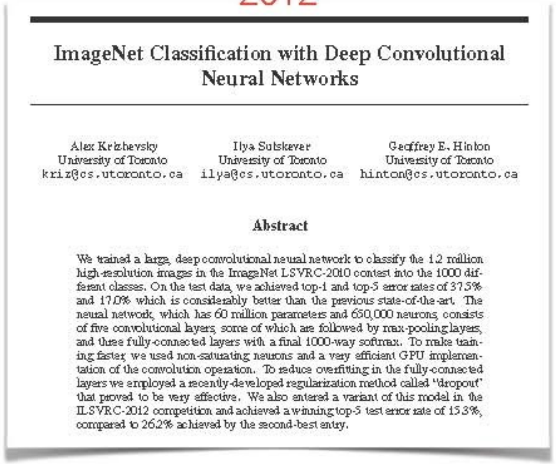


IMAGE CLASSIFICATION



2012



ImageNet Classification with Deep Convolutional Neural Networks

Alex Krizhevsky, University of Toronto, kriz@cs.utoronto.ca
Ilya Sutskever, University of Toronto, ilya@cs.utoronto.ca
Geoffrey E. Hinton, University of Toronto, hinton@cs.utoronto.ca

Abstract

We trained a large, deep convolutional neural network to classify the 1.2 million high-resolution images in the ImageNet LSVRC-2010 contest into the 1000 different classes. On the test data, we achieved top-1 and top-5 error rates of 37.5% and 17.0% which is considerably better than the previous state-of-the-art. The neural network, which has 60 million parameters and 650,000 neurons, consists of five convolutional layers, some of which are followed by max-pooling layers, and three fully-connected layers with a final 1000-way softmax. To make training faster, we used non-saturating neurons and a very efficient GPU implementation of the convolution operation. To reduce overfitting in the fully-connected layers we employed a recently-developed regularization method called "dropout" that proved to be very effective. We also entered a variant of this model in the ILSVRC-2012 competition and achieved a winning top-5 test error rate of 15.3%, compared to 26.2% achieved by the second-best entry.

2015

News & Analysis

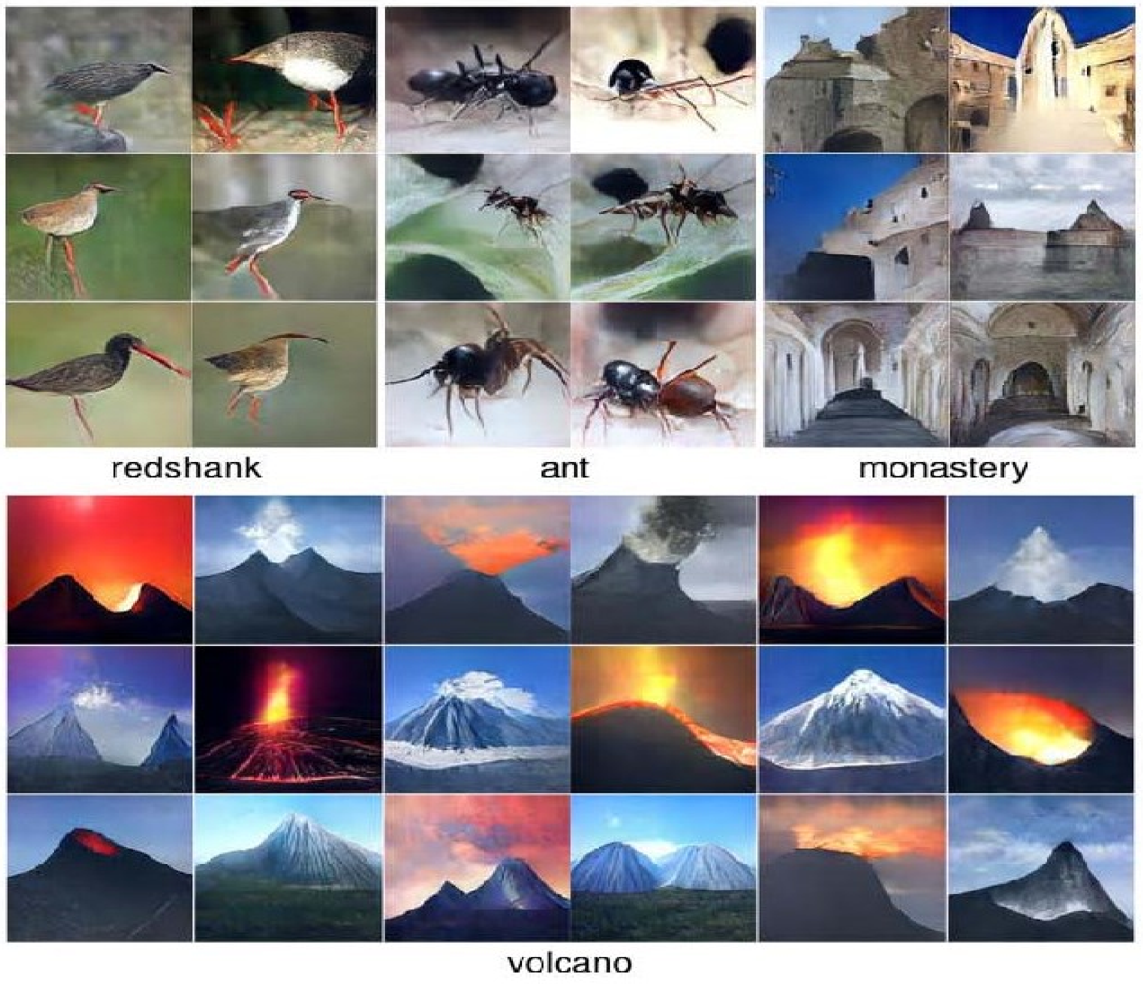
Microsoft, Google Beat Humans at Image Recognition

Deep learning algorithms compete at ImageNet challenge

R. Colin Johnson
2/18/2015 03:15 AM EST
14 comments

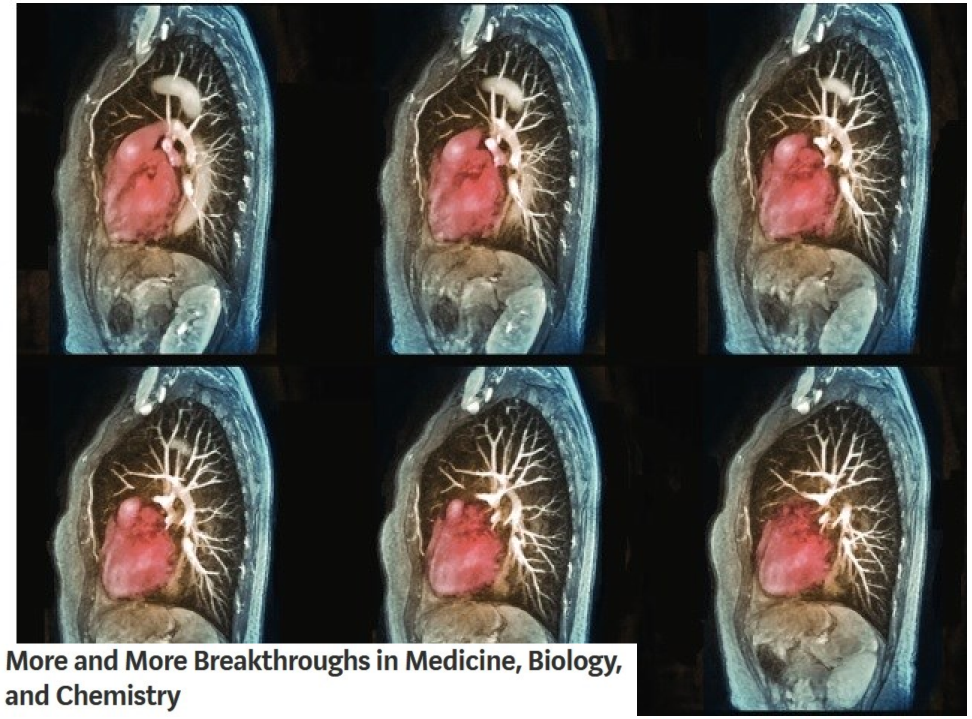
1 saves
LOGIN TO RATE


People Who Don't Exist
A
B
Result of combining A + B


redshank
ant
monastery
volcano

AI Review

- AI is Transforming almost everything including <u>Science and Research</u>



More and More Breakthroughs in Medicine, Biology, and Chemistry

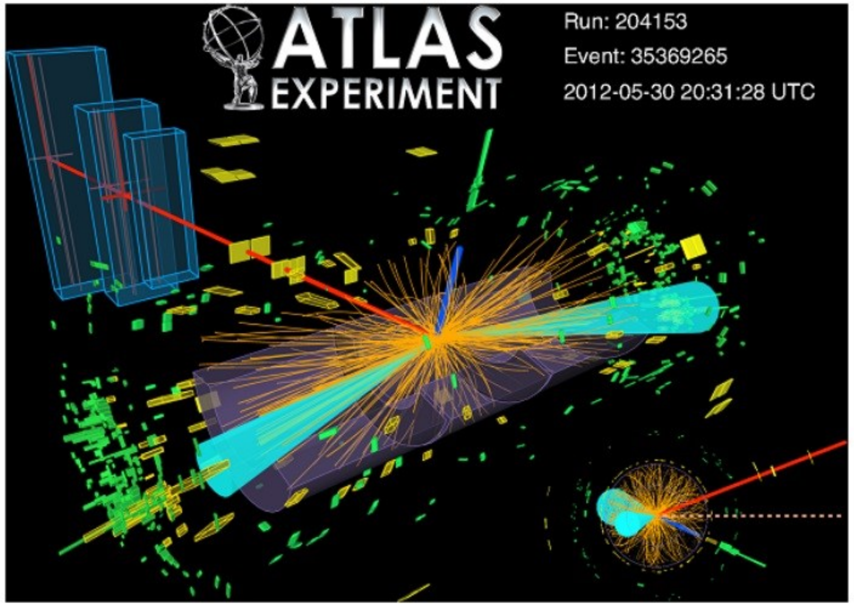
When less is more: Taking MRI scans quickly avoids capturing unwanted movements of the internal organs, but it can result in compromised images. Machine learning, however, could reconstruct improved images based on incomplete data. (Courtesy: Zephyr/Science Photo Library)

Higgs challenge

Higgs Boson Machine Learning Challenge

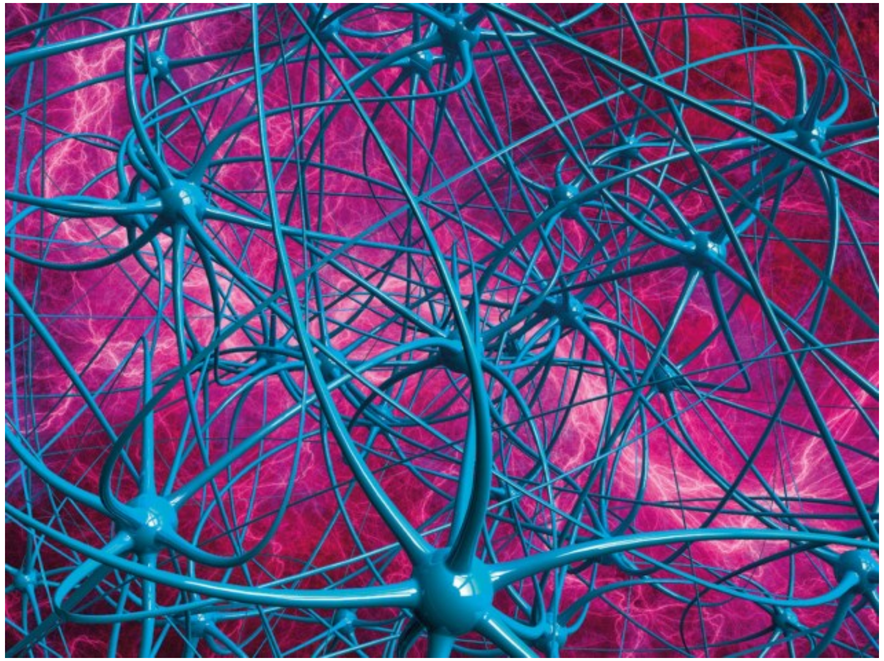
Use the ATLAS experiment to identify the Higgs boson

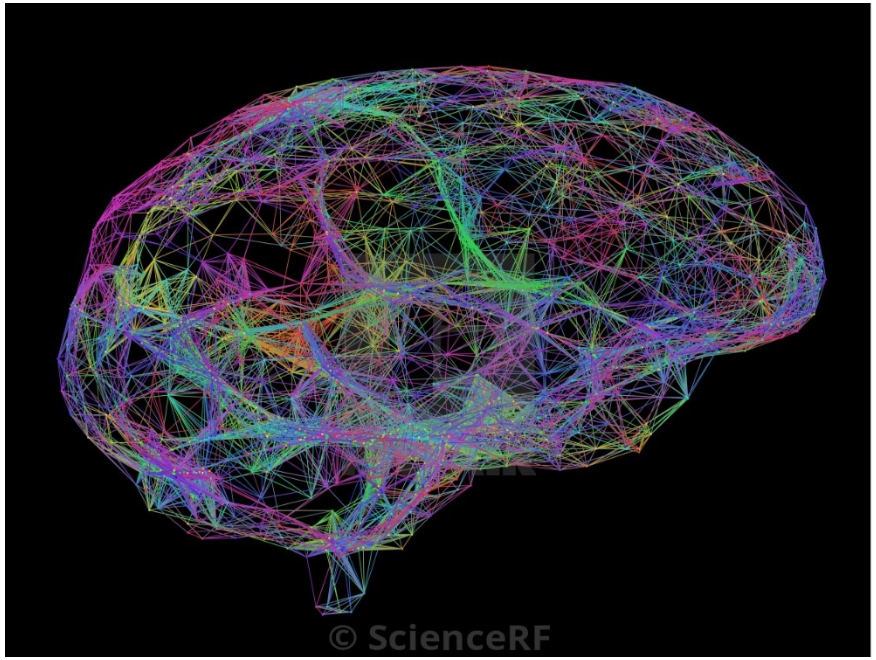
$13,000 · 1,785 teams · 5 years ago

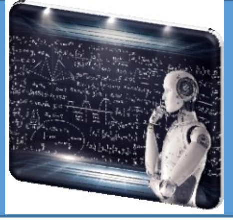


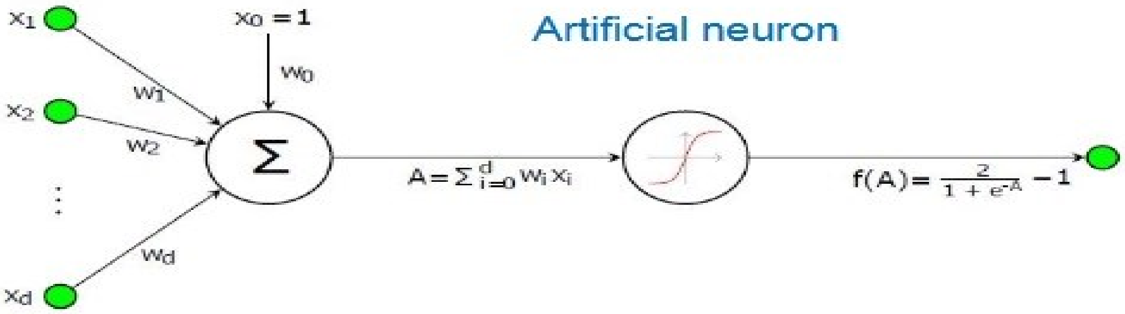
AI Review : ANN

- Artificial Neural Networks : **Simulation of Biological Intelligence**







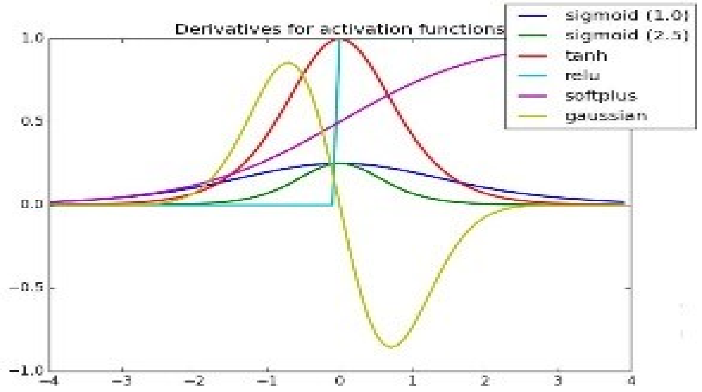


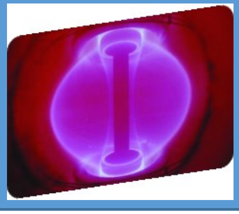
An ANN mimics the behaviour of the biological neuronal networks and consists of an interconnected group of processing elements (referred to as neurons or nodes) arranged in layers.

The first layer, known as the input layer, receives the input variables (x1; x2; …xd). Each connection to the neuron is characterised by a weight (w1; w2; … wd) which can be excitatory (positive weight) or inhibitory (negative weight). Moreover, each layer may have a bias (x0 = 1), which can provide a constant shift to the total neuronal input net activation (A), in this case a sigmoid function:

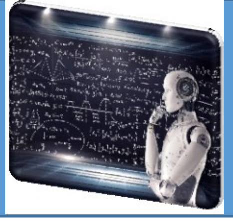
$$A = \sum_{i=1}^{d} w_i x_i + w_0 = \sum_{i=0}^{d} w_i x_i.$$

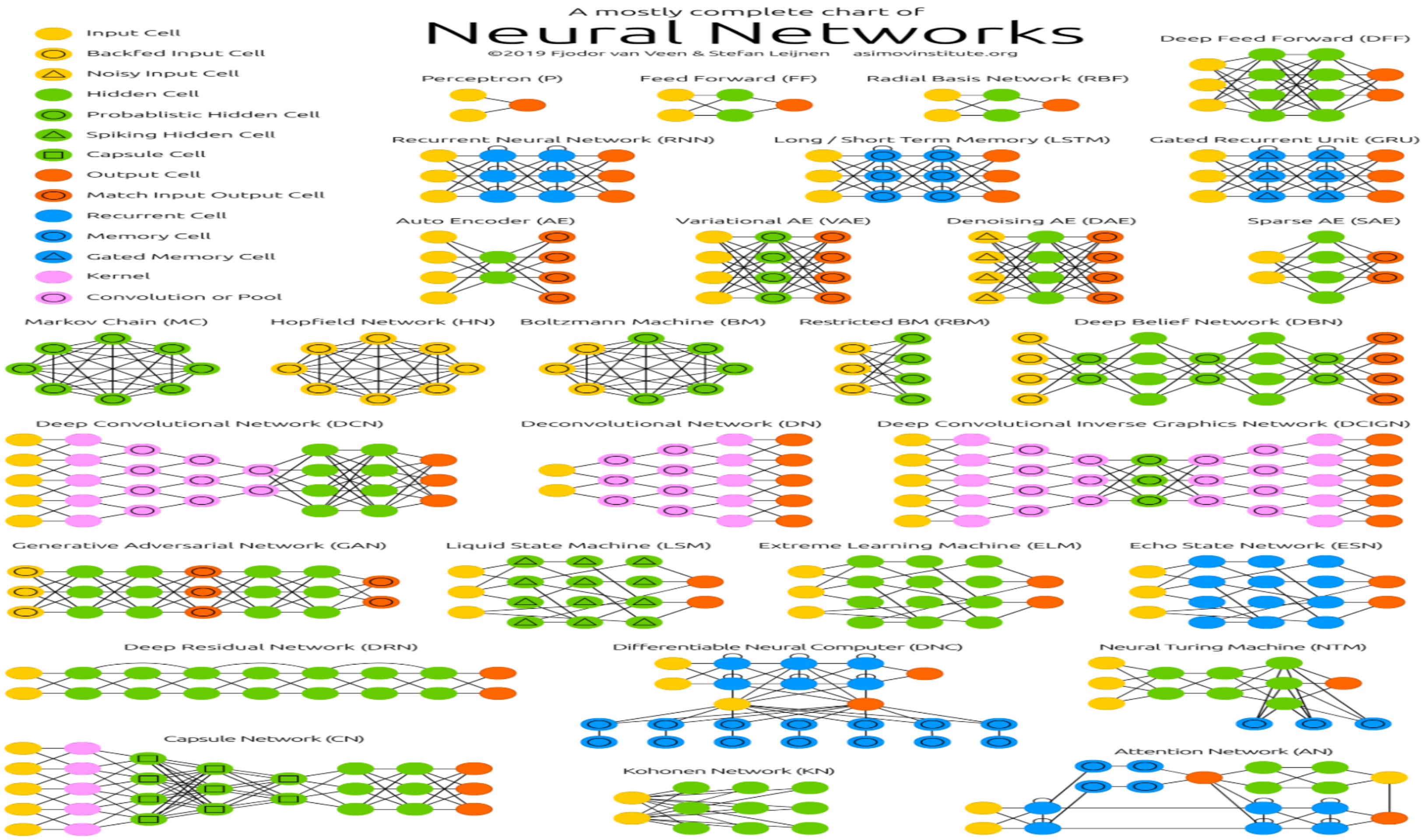
$$f(A) = \frac{2}{1+e^{-A}} - 1,$$





AI Review : ANN



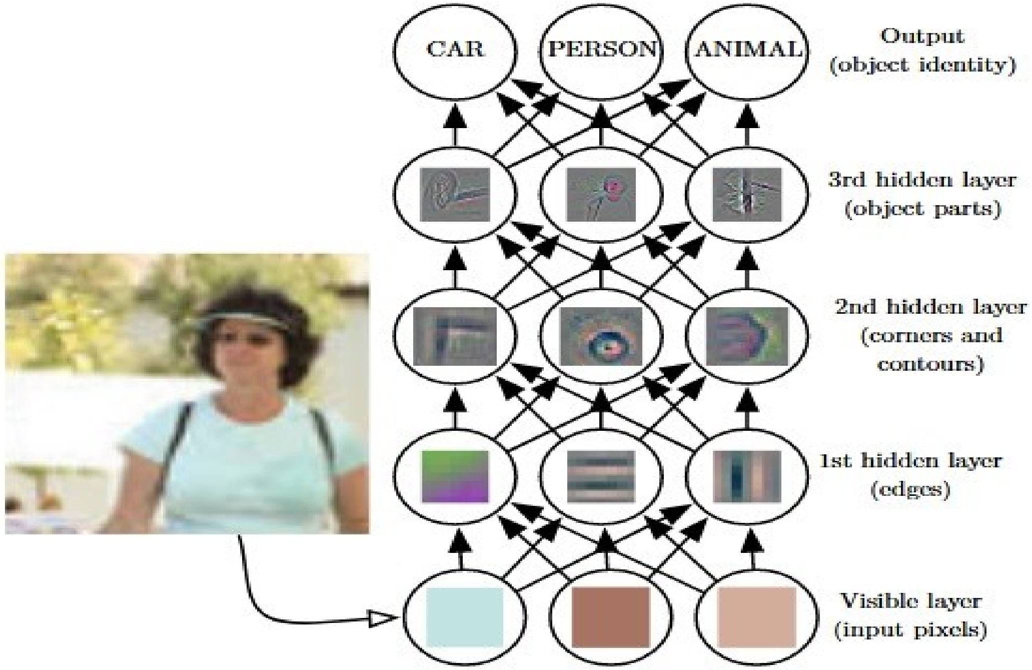


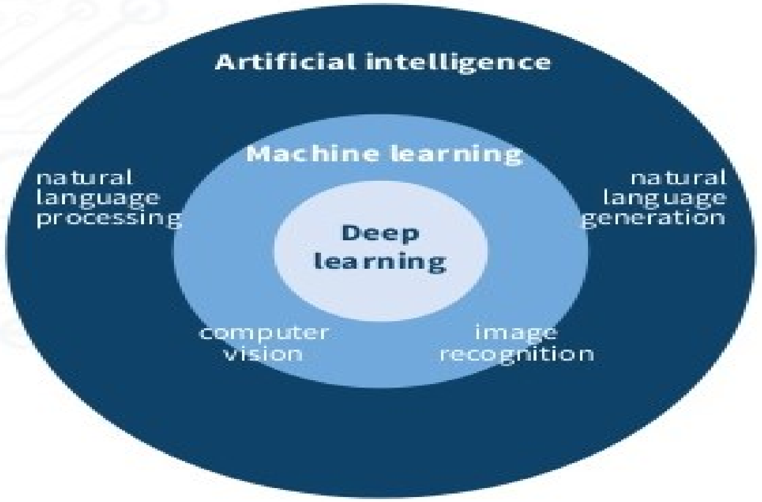
NN architecture: Structure of the networks, and the node connectivity can be adapted for problem at hand

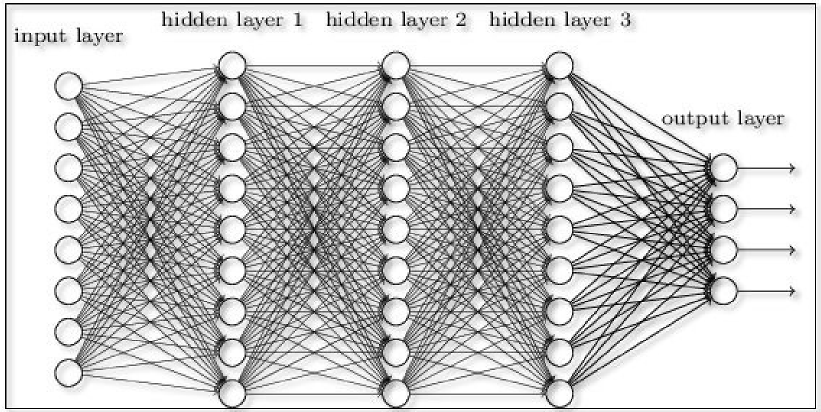
Convolutions: shared weights of neurons, but each neuron only takes subset of inputs

Difficult to train, only recently possible with large datasets, fast computing (GPU) and new training procedures / network structures

Depth: Repeated Composition







Applications in Experimental Research

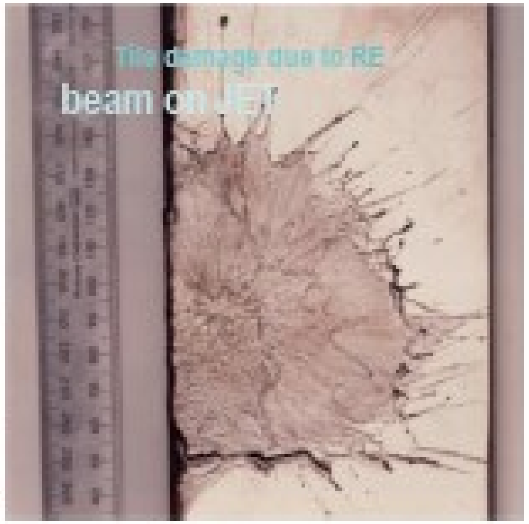
Most Critical Problem for Fusion Energy →

Accurately predict, mitigate, & ideally avoid large-scale major disruptions in magnetically-confined plasmas such as the ITER –the $25B international burning plasma "tokamak"

•Most Effective Approach: Use of big-data-driven statistical/machine-learning predictions guided by observations for the occurrence of disruptions in world-leading facilities such as EUROFUSION "Joint European Torus (JET)" in UK, DIII-D (US), and other tokamaks worldwide such as KSTAR, EAST, JT60-SA (Asia)

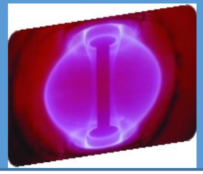
•Recent Status: ~10 years of R&D results (led by JET) using Machine Learning (via Support Vector Machines) on zero-D (scalar) time trace data executed on CPU clusters yielding *success rates in mid-80 to 90% range for JET 30 ms before disruptions,*

BUT > 95% accuracy with false alarm rate < 5% at least 30 milliseconds before actually needed for ITER ! **Reference** – P. DeVries, et al. (2015)

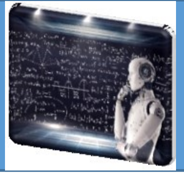


- **Disruption heat fluxes can *reduce component lifetime***
  (e.g. divertor target ablation)

- **Damage to in-vessel components *can require shutdown for repair***



Applications in Experimental Research



AI/DL/Machine Learning Workflow

Identify Signals
- **Classifiers**

Preprocessing and feature extraction

Normalization

Train model, Hyper parameter tuning

Use model for prediction

Princeton U/PPPL AI/DL predictions now works with *multi-D time trace signals* (beyond zero-D)

Measured sequential data arranged in patches of equal length for training

All data placed on appropriate numerical scale ~ O(1) e.g., Data-based with all signals divided by their standard deviation

- All available data analyzed;
- Train LSTM (Long Short Term Memory Network) iteratively;
- Evaluate using ROC (Receiver Operating Characteristics) and cross-validation loss for every epoch (equivalent of entire data set for each iteration)

Apply AI/DL/ML software on plasma state signals

Artificial Intelligence/Deep Learning brings new technology to accelerate progress

"Predicting Disruptive Instabilities in Controlled Fusion Plasmas through Deep Learning"

NATURE: (accepted for publication, Jan. 2019, published, April 17, 2019 –
DOI: 10.1038/s41586-019-1116-4)

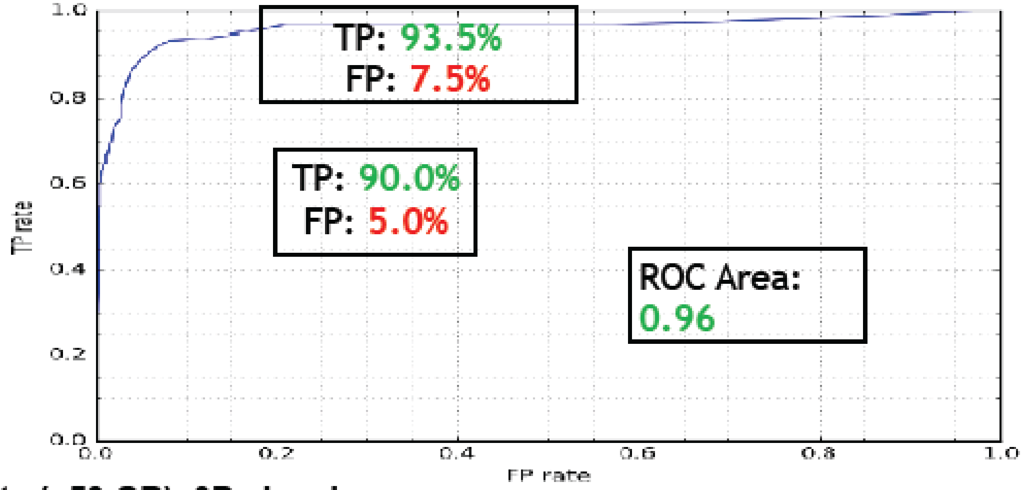
Princeton's Fusion Recurrent Neural Network code (FRNN) uses convolutional & recurrent neural network components to integrate both spatial and temporal information for predicting disruptions in tokamak plasmas with unprecedented accuracy and speed on top supercomputers

FRNN Code PERFORMANCE: ROC CURVES

JET ITER-like Wall Cases @30ms before Disruption

EUROfusion

Performance Tradeoff: Tune True Positives (good: correctly caught disruption) **vs. False Positives** (bad: safe shot incorrectly labeled disruptive).

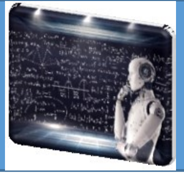


JET Data (~50 GB), 0D signals:

- Training: on **4100 shots** from **JET C-Wall** campaigns
- Testing **1200 shots** from **Jet ILW** campaigns
- **All shots used**, no signal filtering or removal of shots

JET Data courtesy of J. Vega and A. Murari

Applications in Experimental Research



PHYSICS OF PLASMAS **25**, 080901 (2018)

Deep learning: A guide for practitioners in the physical sciences

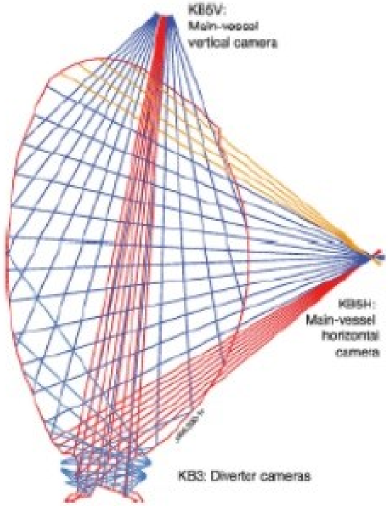
Brian K. Spears,[1,a),b)] James Brase,[1] Peer-Timo Bremer,[1] Barry Chen,[1] John Field,[1] Jim Gaffney,[1] Michael Kruse,[1] Steve Langer,[1] Katie Lewis,[1] Ryan Nora,[1] Jayson Luc Peterson,[1] Jayaraman Jayaraman Thiagarajan,[1] Brian Van Essen,[1] and Kelli Humbird[2]

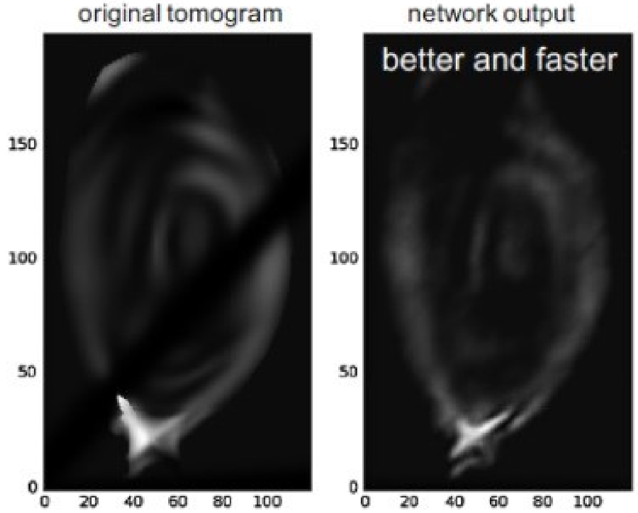
TABLE I. Summary of example applications, their inputs, and their outputs.

Discipline	Example application	Input features	Output predictions
Plasma physics	Magnetic confinement fusion[2,24,26]	Plasma and device measurements	Device state (disrupted, non-disrupted, etc.)
Plasma physics	Closure of turbulence models[25]	Analytical source terms	Rapid, approximate source terms
Plasma physics	Inertial confinement fusion[13,20,23]	Simulated implosion input parameters (laser brightness, target dimensions, etc.)	Returns simulated implosion observations (neutron yield, ion temperature, etc.)
Astronomy and astrophysics	Galaxy classification[11]	Galaxy images	Identifies various phases of blue nugget galaxies
Astronomy and astrophysics	Orbital stability[16]	Orbital parameters (e_{bin}, μ, a_p/a_{bin}, ϵ)	Predicts stability or instability of circumbinary orbits
Biology	Cellular analysis[27]	Neuron images	Identifies dead neurons in mixed living/dead populations
Biology	Drug discovery[3]	Description of a target protein structure	Suggests molecular binder candidates for protein interaction
Particle physics	Particle detection[1]	Particle momentum and direction	Particle type (lepton, quark jet, etc.)

Deep Learning for real-time plasma control

Plasma tomography: Use CNNs to reconstruct cross-section from projections





Traditional inversion schemes:

- Varying runtime
- Dependence on additional data (magnetic equilibrium)
- Not real-time capable

Deep learning for plasma tomography using the bolometer system at JET

PhD student at IPP

Francisco A. Matos[a], Diogo R. Ferreira[a,*], Pedro J. Carvalho[b], JET Contributors[1]

Deep Learning for Plasma Tomography

Francisco Duarte Pinto de Almeida Matos

Thesis to obtain the Master of Science Degree in

Information Systems and Computer Engineering

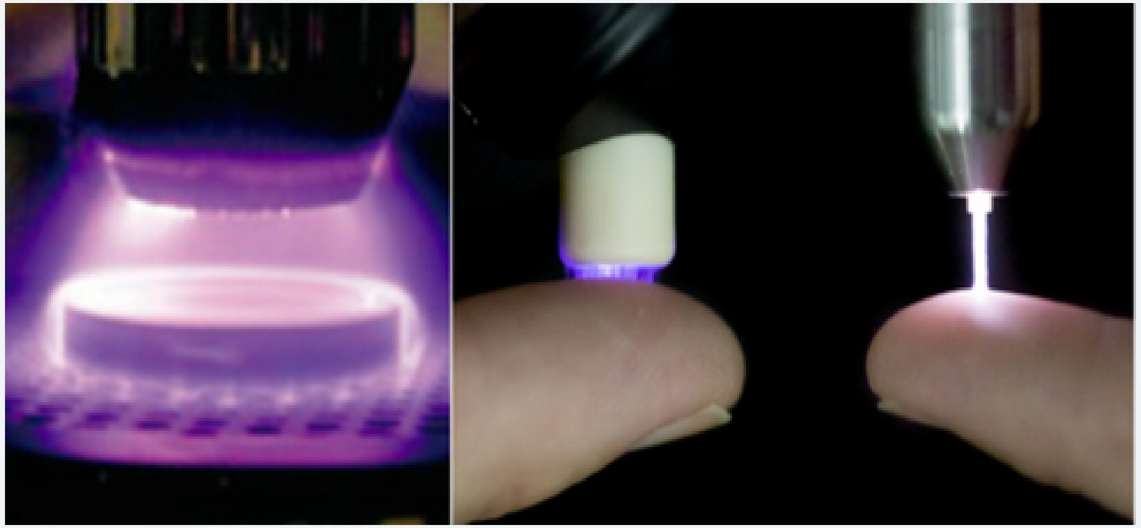
Supervisor(s): Prof. Diogo Manuel Ribeiro Ferreira
Dr. Pedro Jorge de Paula Carvalho

Examination Committee

Chairperson: Prof. José Carlos Martins Delgado
Supervisor: Prof. Diogo Manuel Ribeiro Ferreira
Member of the Committee: Prof. Andreas Miroslaus Wichert

June 2016

Machine Learning and Predictive Control for Non-Equilibrium Plasma Treatment of Complex Surfaces

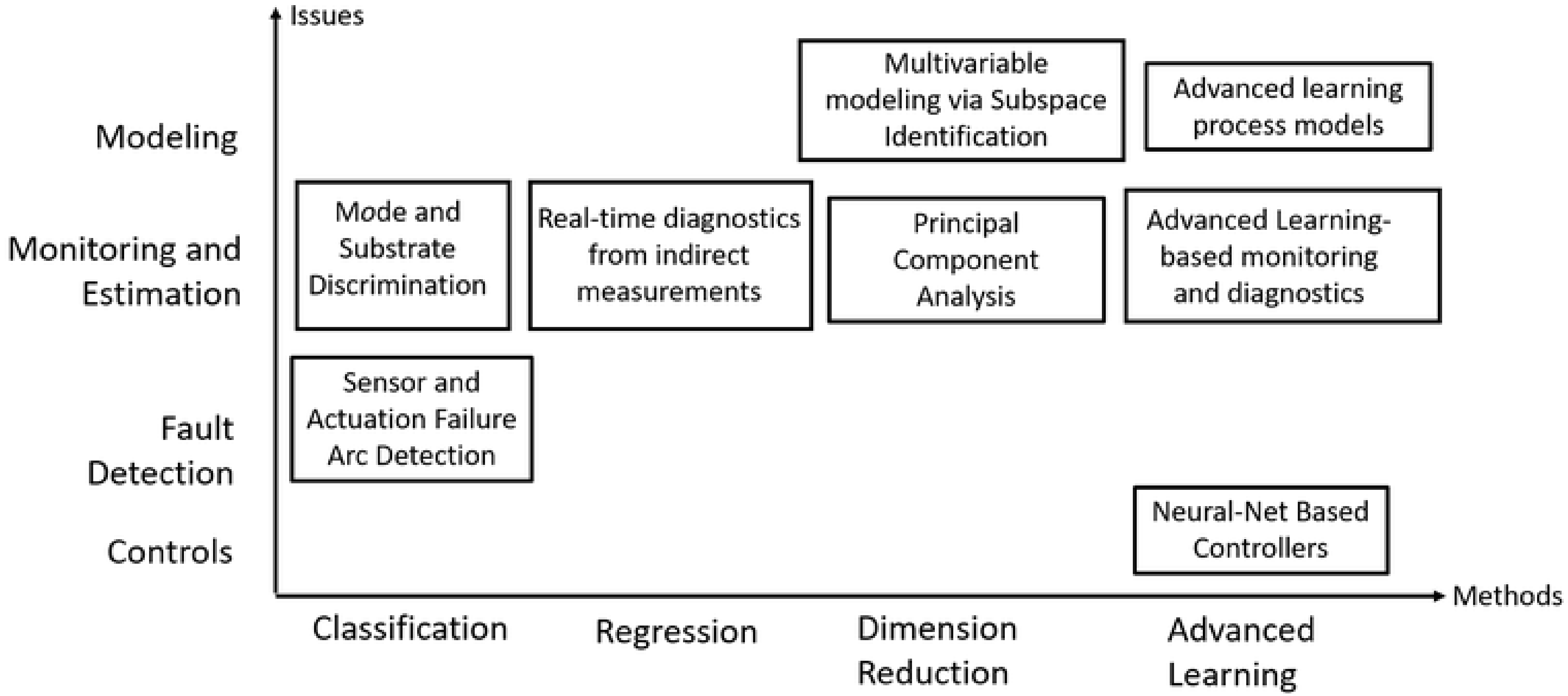


Non-equilibrium plasmas for surface functionalization of bio-materials and plasma medicine.

Lack of mechanistic insights into interactions of non-equilibrium plasmas (NEPs) with complex surfaces (e.g., in plasma medicine, plasma catalysis, or plasma bioprocessing), poses a major challenge in reliable and effective NEP processing. For example, therapeutic applications of NEPs in plasma medicine demand selective and high-efficacy treatment of complex biological substrates, while ensuring safe and reproducible treatment. This research investigates the application of machine learning and predictive control for characterizing and controlling the underpinning mechanisms of plasma-interfacial reactions with complex surfaces. The research is conducted in collaboration with the **Graves Lab**.

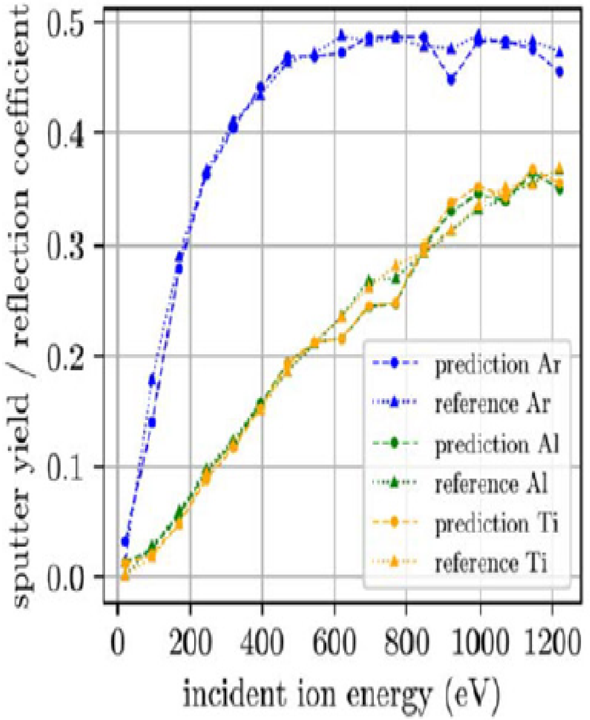
Applications in Theoretical Research

Plasma Machine Learning



Applications in Theoretical Research

J. Phys. D: Appl. Phys. **52** (2019) 30LT02 Letters



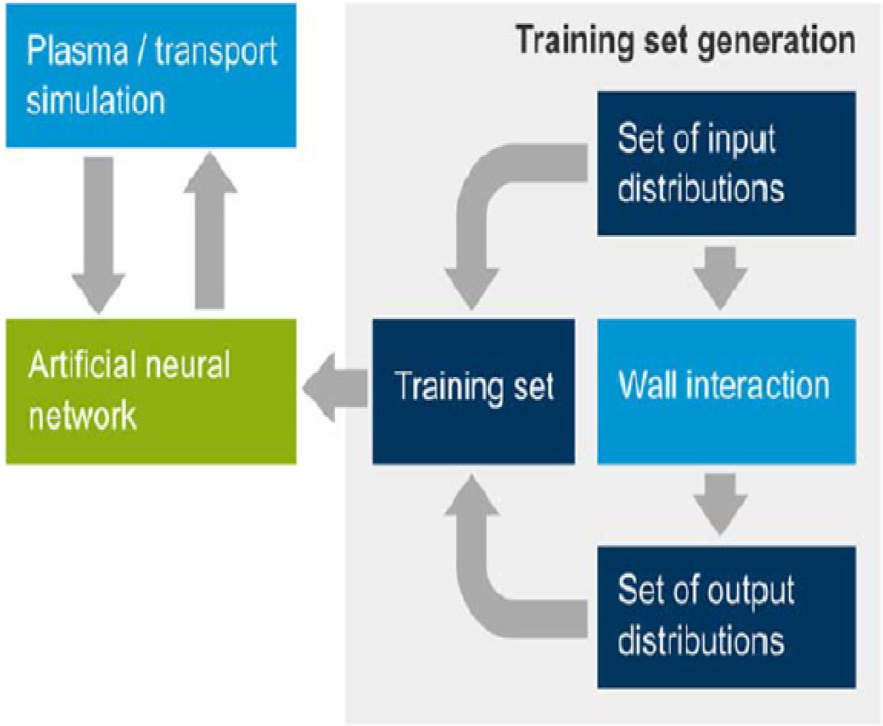
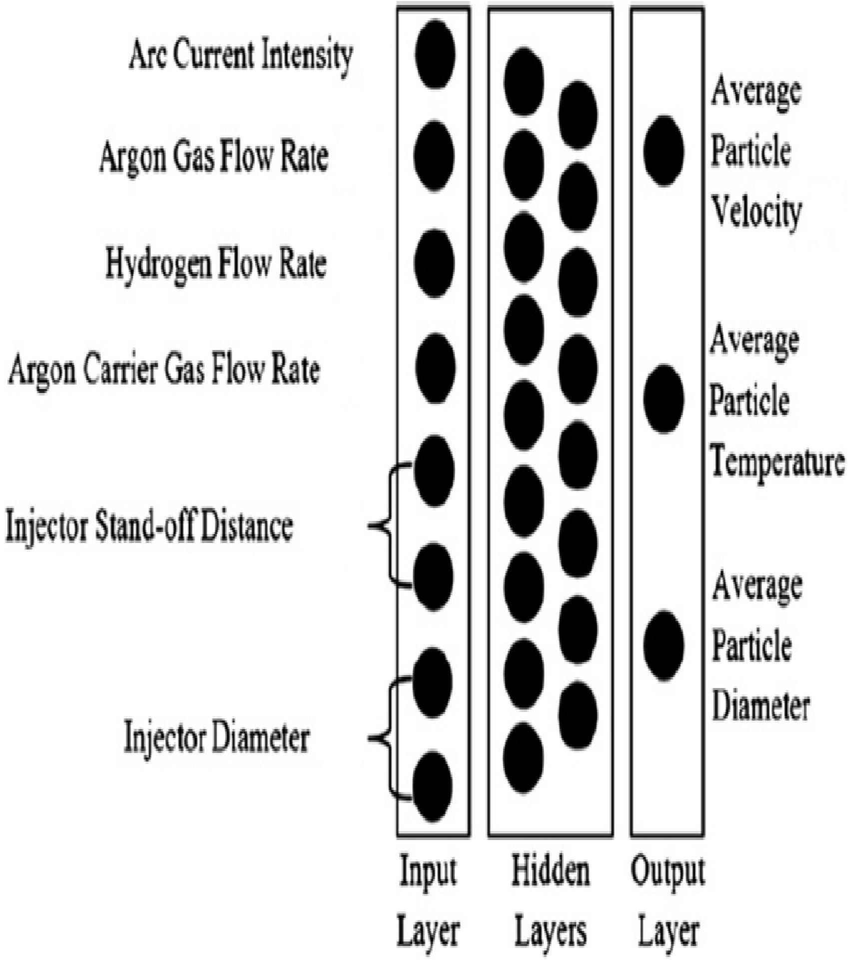
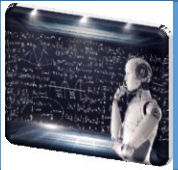


Figure 1. Supervised learning enables construction of computationally efficient surrogate models from theoretical simulation data for multiscale modeling of plasma-surface interactions across multiple length- and time-scales. Here, artificial neural networks were used to develop a plasma–surface interface model for a plasma sputtering process. The interface model was used for predicting the energetic and angular distribution of surface species ejected into the plasma as a function of energy distributions of incident species. Accordingly, the



Applications in Theoretical Research



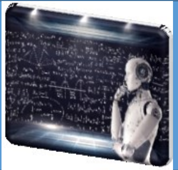
J. Phys. D: Appl. Phys. **52** (2019) 30LT02

Letters

Table 1. An overview of potential applications of ML for modeling, diagnostics, and control of NEPs.

	Supervised learning (e.g. regression, neural networks, kriging, support vector machines)	Unsupervised learning (e.g. clustering, dimension reduction)	Reinforcement learning
Predictive modeling	Learning nonlinear mappings for plasma-surface interactions [4, 22], learning inexpensive surrogate models from theoretical simulation data [7], plasma dose quantification	Selection of relevant input features for building simpler models from data [5]	
Diagnostics	Inference of plasma and surface properties from spectral data [6, 10, 32],	Extraction of latent information from measurements [46–48],	
	Detection of abnormal drifts and variabilities [6]		
Process control	Learning multivariable input–output mappings of process dynamics for model-based control [45, 56, 57]		Learning-based control

Applications in Theory : Numerical Simulation



Turbulent transport in toroidal magnetic confinement devices, such as tokamaks, is one of the limiting factors for achieving viable fusion energy. Reactor design and plasma scenario optimisation demands both accurate and tractable predictive turbulence calculations. Neural network surrogate physics models provides a pathway to this goal. Databases of reduced order turbulence model output act as training sets for neural network regression. A key aspect is the customisation of output regression variables and optimisation cost functions in a physics-informed manner. The resultant surrogate model is 1012 times faster than direct numerical simulation and 106 times faster than the reduced order model itself.

Jonathan Citrin leads a computational physics research group at the Dutch Institute for Fundamental Energy Research (Differ) and his focus is on plasma simulation.

Surrogate modelling using feed forward neural networks for turbulent transport in fusion plasmas

1∗van de Plassche, K.L.,1Citrin, J.,2Bourdelle, C.,3Camenen, Y.,4Casson, F.J.,1Dagnelie, V.I.,**5 Felici, F.,**1Ho, A&†JET Contributors

- Surrogate models approximate more complex, computationally expensive models while being faster to run. They can be used for tasks where using the slow original model is in-feasible, for example in optimization and control. In this work we apply this methodology within a fusion energy context, using feed-forward neural networks (FFNNs) as a surro-gate model five orders of magnitude faster than the underlying model: the quasilinear turbulent transport code QuaLiKiz[1, 2]. QuaLiKiz is used to describe heat, particle, and momentum transport in tokamaks, and was used to create a large database of3.108fluxcalculations using 1.3 MCPUh on HPC resources (Edison@NERSC). Embedding known physical constraints in the training of the networks is essential for the surrogate model to perform well in transport predictions. As such, we show the importance of choosing theright cost function and more fundamentally, choosing which target variables the networks have to be trained on. Custom figures of merit and visualization tools were developed to aid with neural network accuracy verification . The neural network surrogate turbulent transport model is applied within the **Rapid Plasma Transport Simulator RAPTOR**[3, 4] and the integrated modelling suite JINTRAC[5] to predict the temperature and density evolution of JET fusion plasmas, in excellent agreement with the original QuaLiKiz model, **yet orders of magnitude faster**. This allows us to simulate one second of plasma evolution in 10 seconds, a speed that is unprecedented for first-principle based transport simulations, opening up new avenues for tokamak scenario optimization and real time control applications.

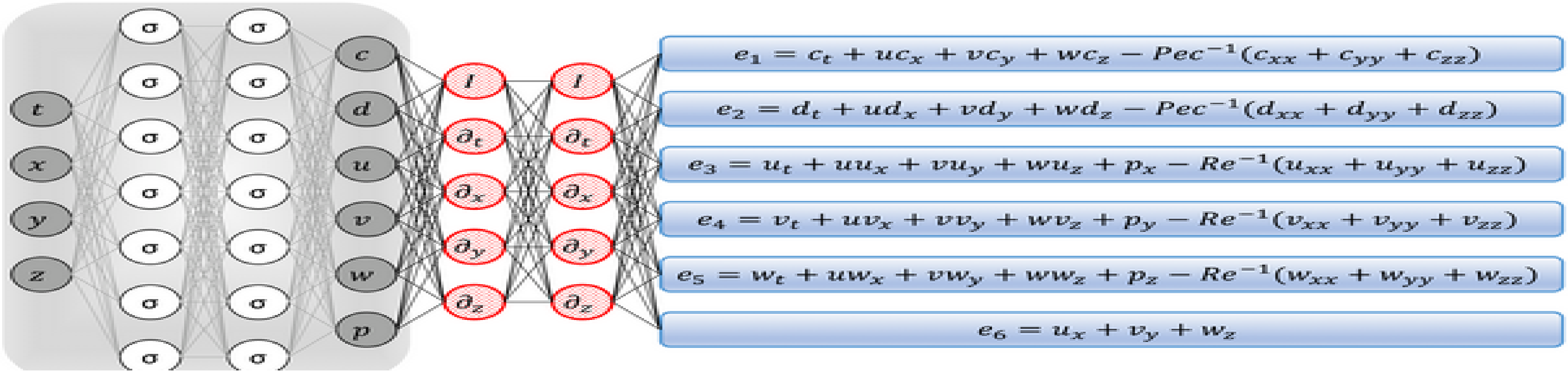
Applications in Theory : Differential Equations

- A **Navier-Stokes** Informed Deep Learning Framework for Assimilating Flow : **Fluid Mechanics**

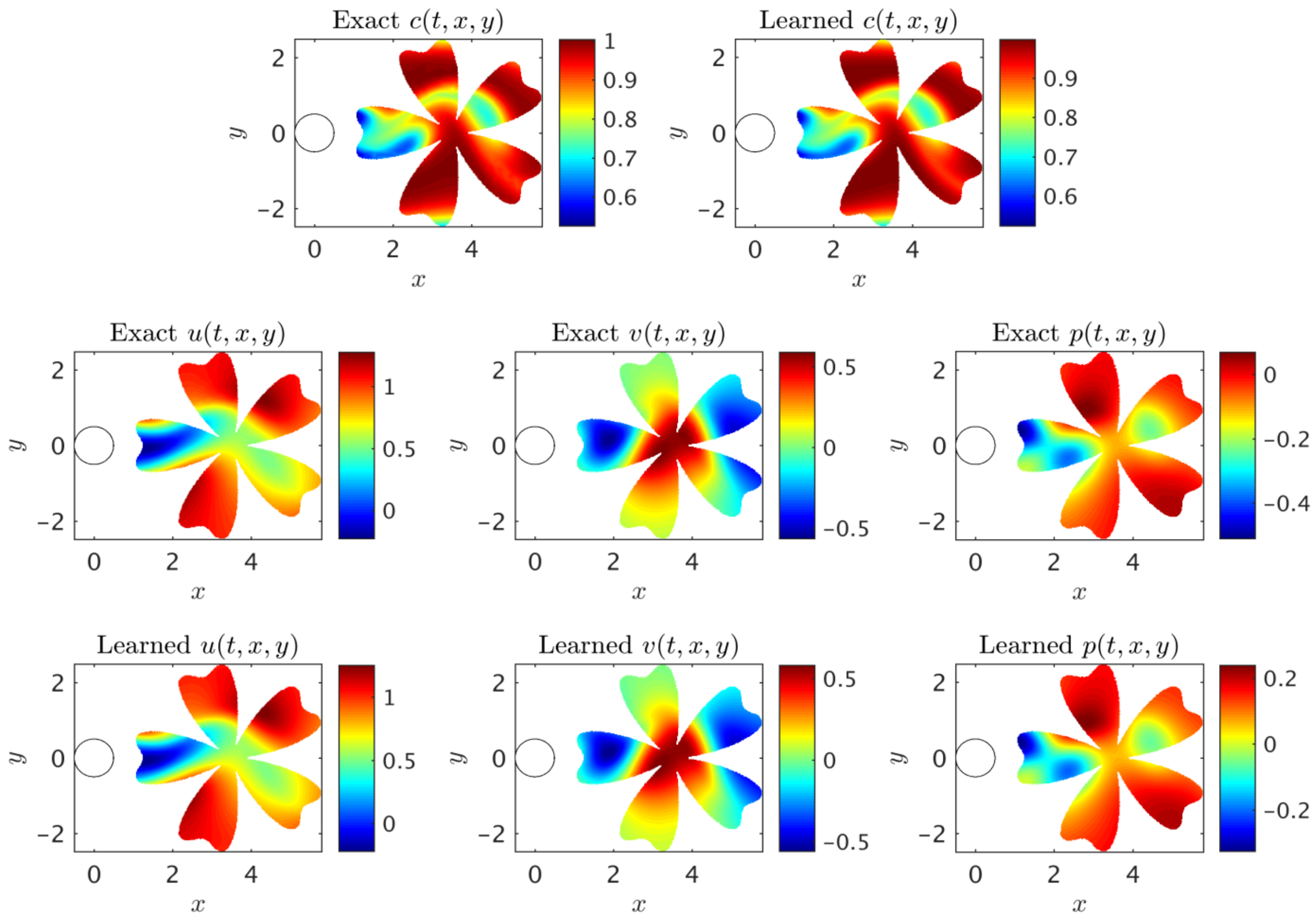
$$(t, x, y, z) \longmapsto (c, d, u, v, w, p)$$

by a deep neural network and obtain the following *Navier-Stokes informed neural networks* (see the following figure) corresponding to the above equations; i.e.,

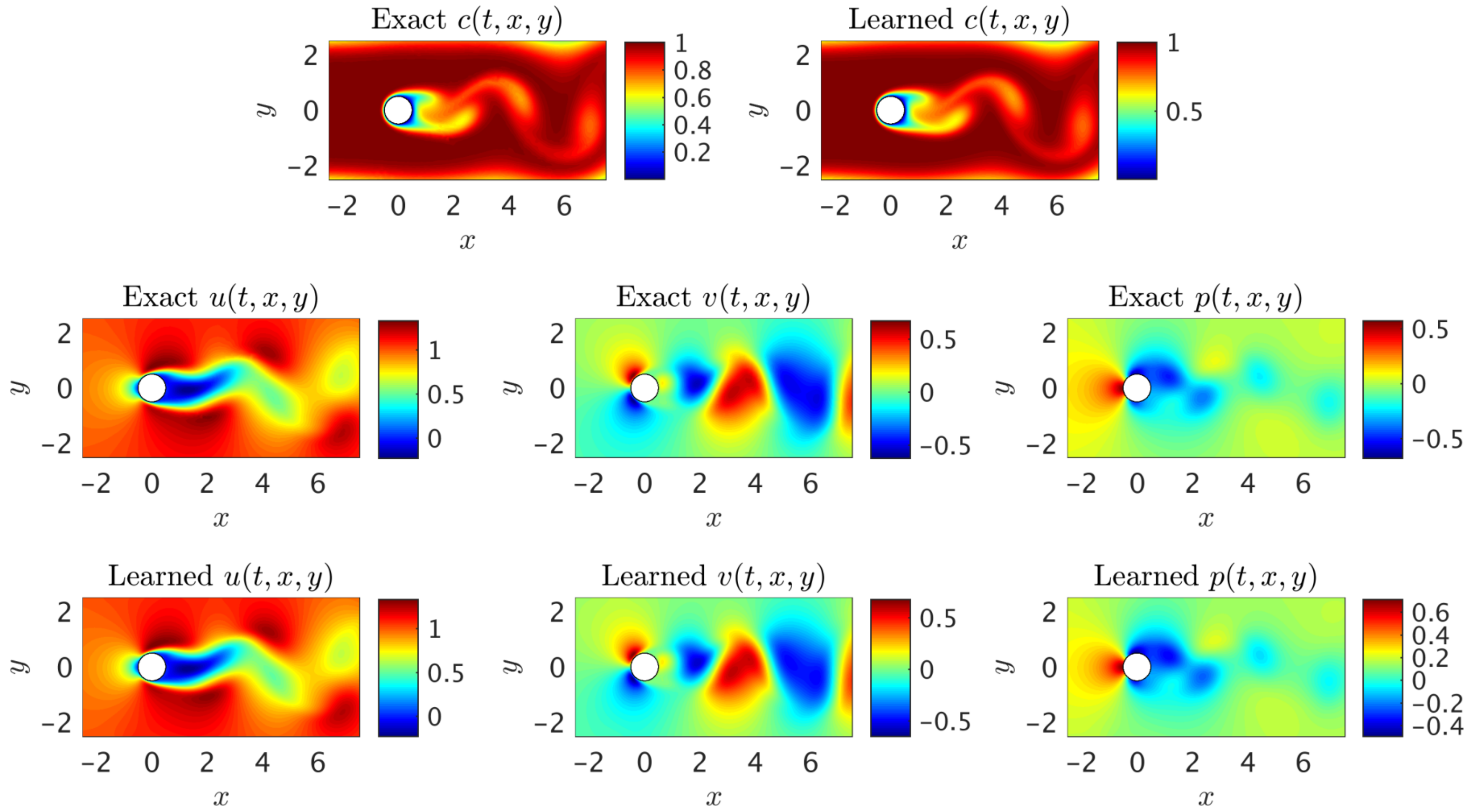
$$\begin{aligned}
e_1 &:= c_t + uc_x + vc_y + wc_z - Pec^{-1}(c_{xx} + c_{yy} + c_{zz}),\\
e_2 &:= d_t + ud_x + vd_y + wd_z - Pec^{-1}(d_{xx} + d_{yy} + d_{zz}),\\
e_3 &:= u_t + uu_x + vu_y + wu_z + p_x - Re^{-1}(u_{xx} + u_{yy} + u_{zz}),\\
e_4 &:= v_t + uv_x + vv_y + wv_z + p_y - Re^{-1}(v_{xx} + v_{yy} + v_{zz}),\\
e_5 &:= w_t + uw_x + vw_y + ww_z + p_z - Re^{-1}(w_{xx} + w_{yy} + w_{zz}),\\
e_6 &:= u_x + v_y + w_z.
\end{aligned}$$



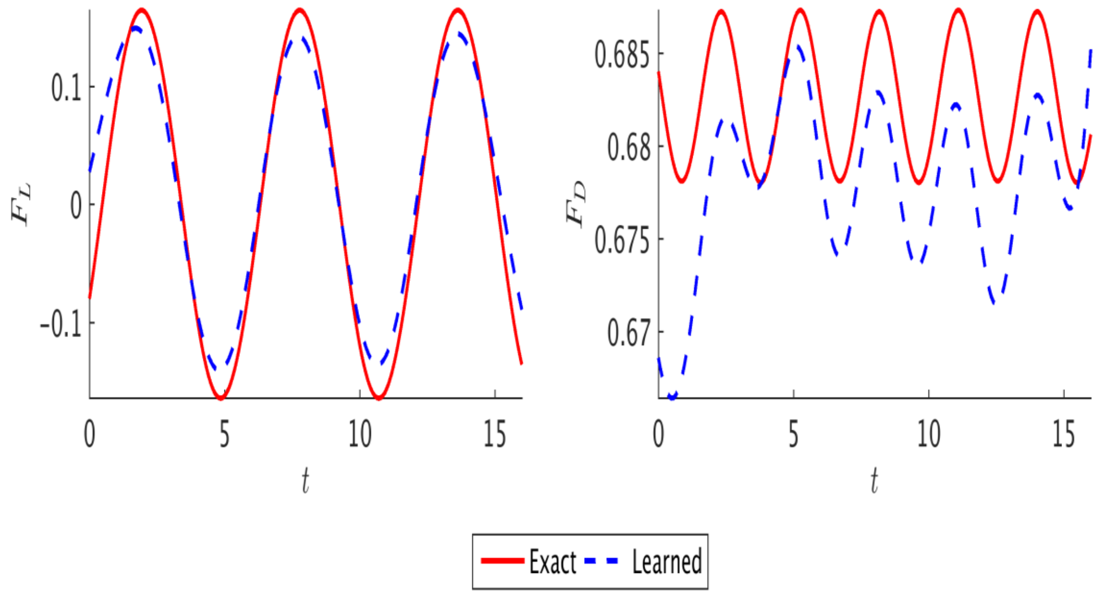
Applications in Theory : Differential Equations

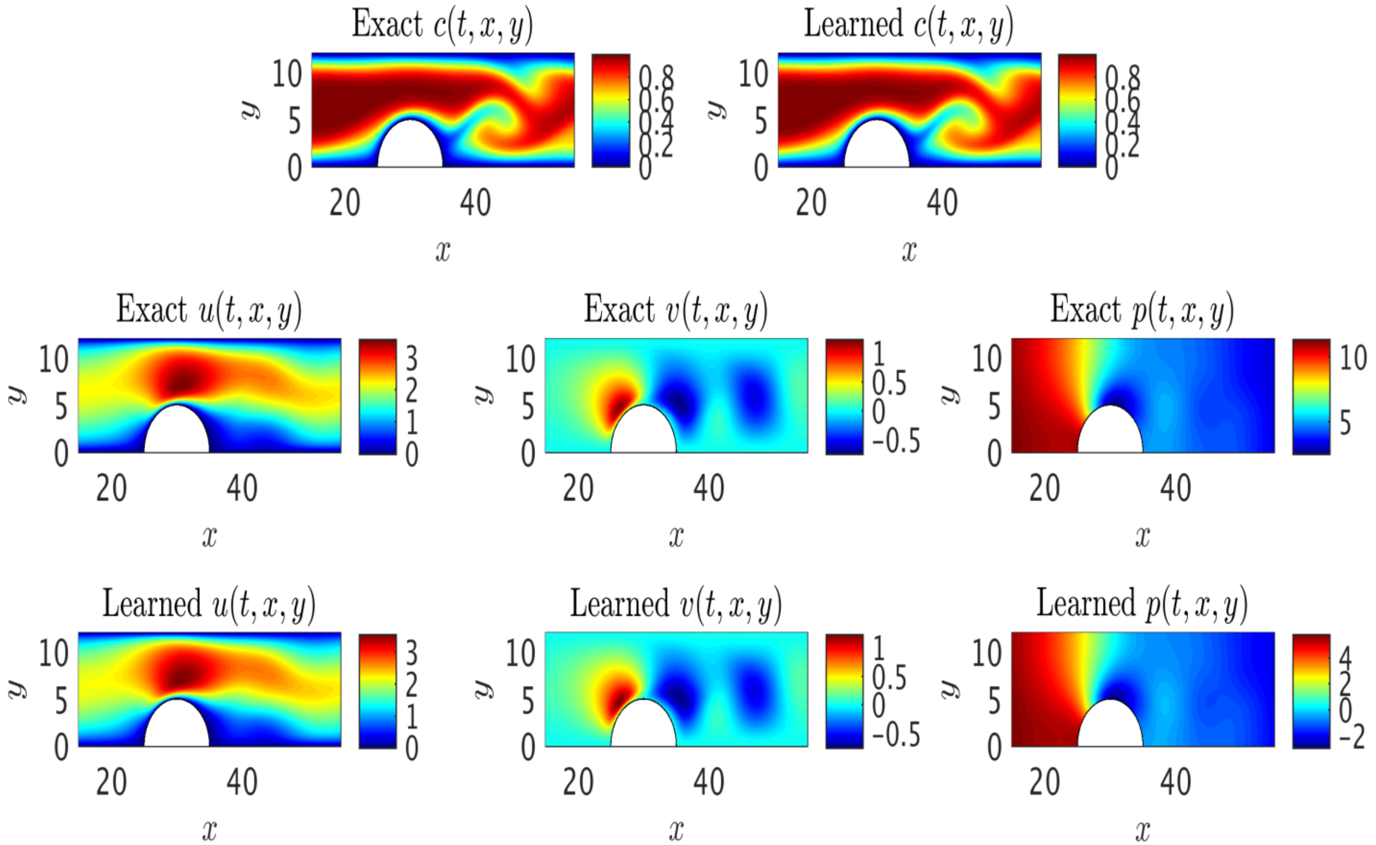


2D Flow past a circular cylinder: A representative snapshot of the input data on the concentration of the passive scalar is shown in the top left panel of this figure, where on the right the same concentration field is reconstructed based on the predictions of our algorithm.

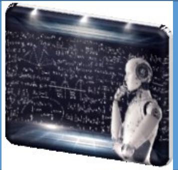

Exact $c(t,x,y)$
Learned $c(t,x,y)$
Exact $u(t,x,y)$
Exact $v(t,x,y)$
Exact $p(t,x,y)$
Learned $u(t,x,y)$
Learned $v(t,x,y)$
Learned $p(t,x,y)$
x
y

Applications in Theory : Differential Equations





Applications in Theory : Differential Equations



Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations

Authors	Maziar Raissi, Paris Perdikaris, George E Karniadakis
Publication date	2019/2/1
Journal	Journal of Computational Physics
Volume	378
Pages	686-707
Publisher	Academic Press
Description	We introduce *physics-informed neural networks* – neural networks that are trained to solve supervised learning tasks while respecting any given laws of physics described by general nonlinear partial differential equations. In this work, we present our developments in the context of solving two main classes of problems: data-driven solution and data-driven discovery of partial differential equations. Depending on the nature and arrangement of the available data, we devise two distinct types of algorithms, namely continuous time and discrete time models. The first type of models forms a new family of *data-efficient* spatio-temporal function approximators, while the latter type allows the use of arbitrarily accurate implicit Runge–Kutta time stepping schemes with unlimited number of stages. The effectiveness of the proposed framework is demonstrated through a collection of classical problems in fluids, quantum ...
Total citations	Cited by 172

Applications in Theory : Differential Equations

692 M. Raissi et al. / Journal of Computational Physics 378 (2019) 686–707

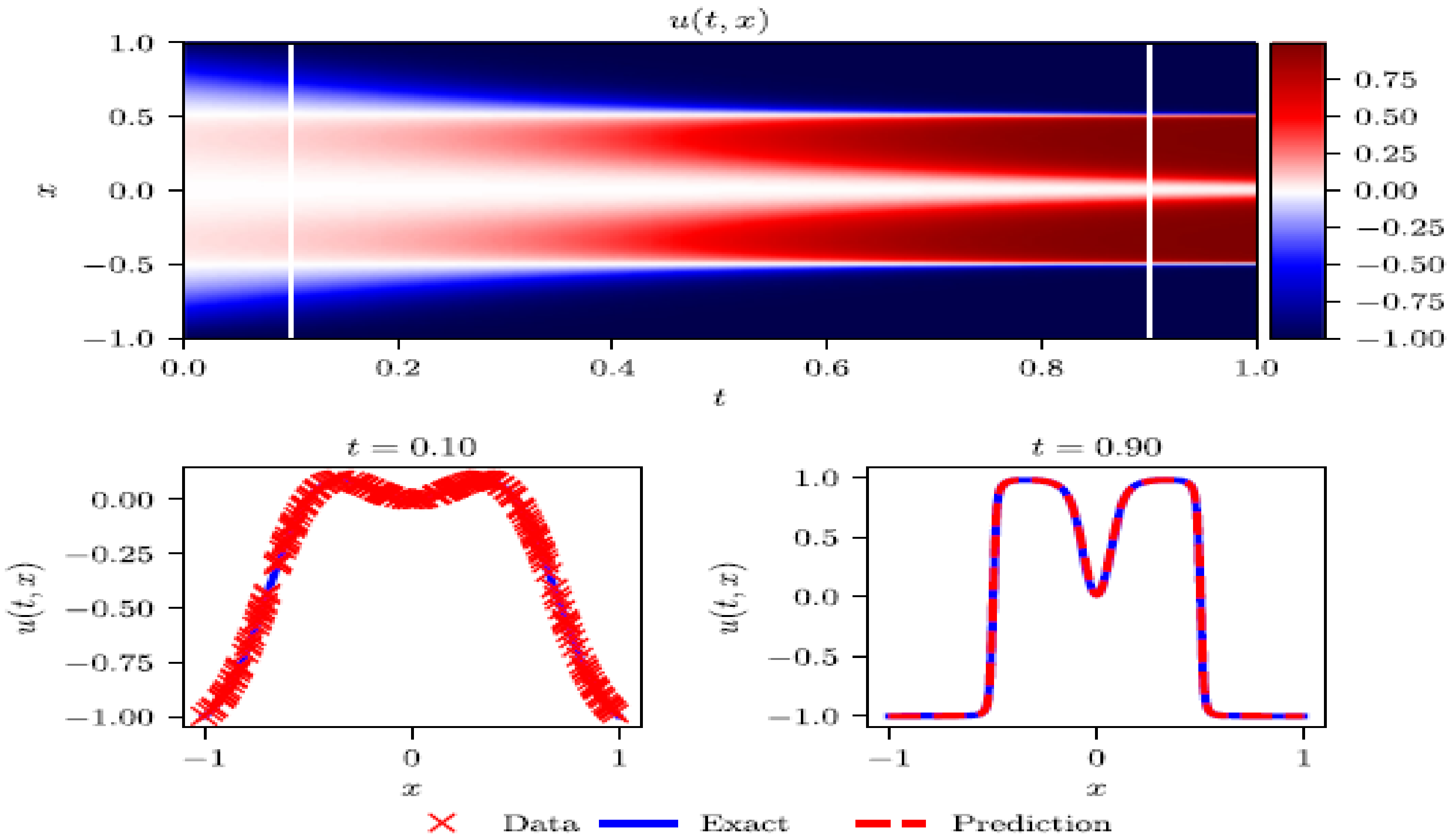
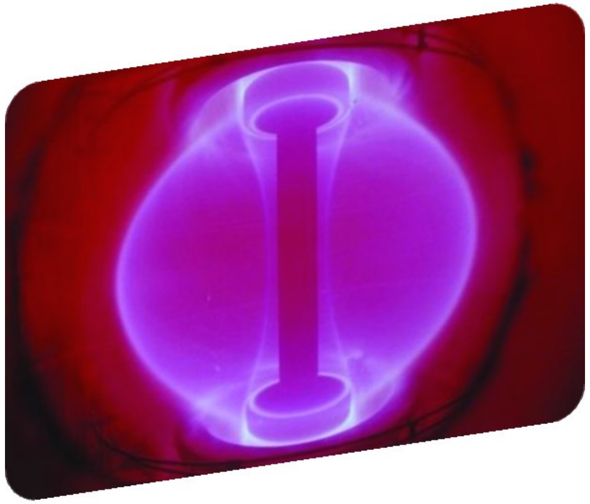
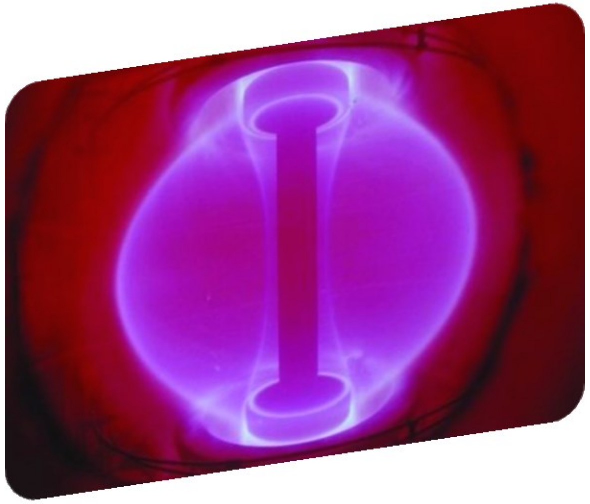


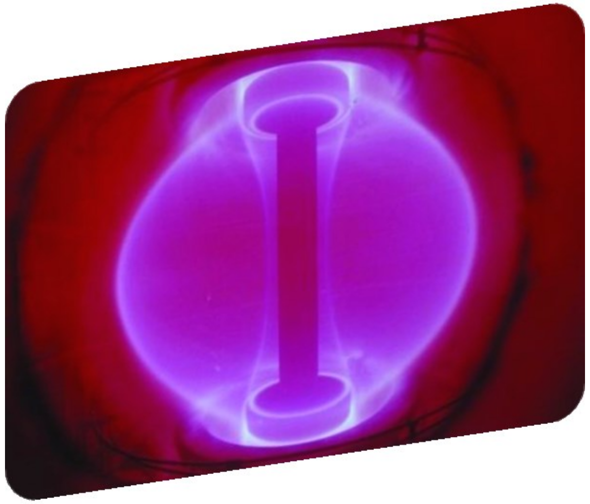
Fig. 2. *Allen–Cahn equation: Top:* Solution $u(t,x)$ along with the location of the initial training snapshot at $t = 0.1$ and the final prediction snapshot at $t = 0.9$. *Bottom:* Initial training data and final prediction at the snapshots depicted by the white vertical lines in the top panel. The relative $\mathbb{L}_2$ error for this case is $6.99 \cdot 10^{-3}$. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

Conclusion : Intelligent Computations



Questions & Answers : Discussion





Thank you