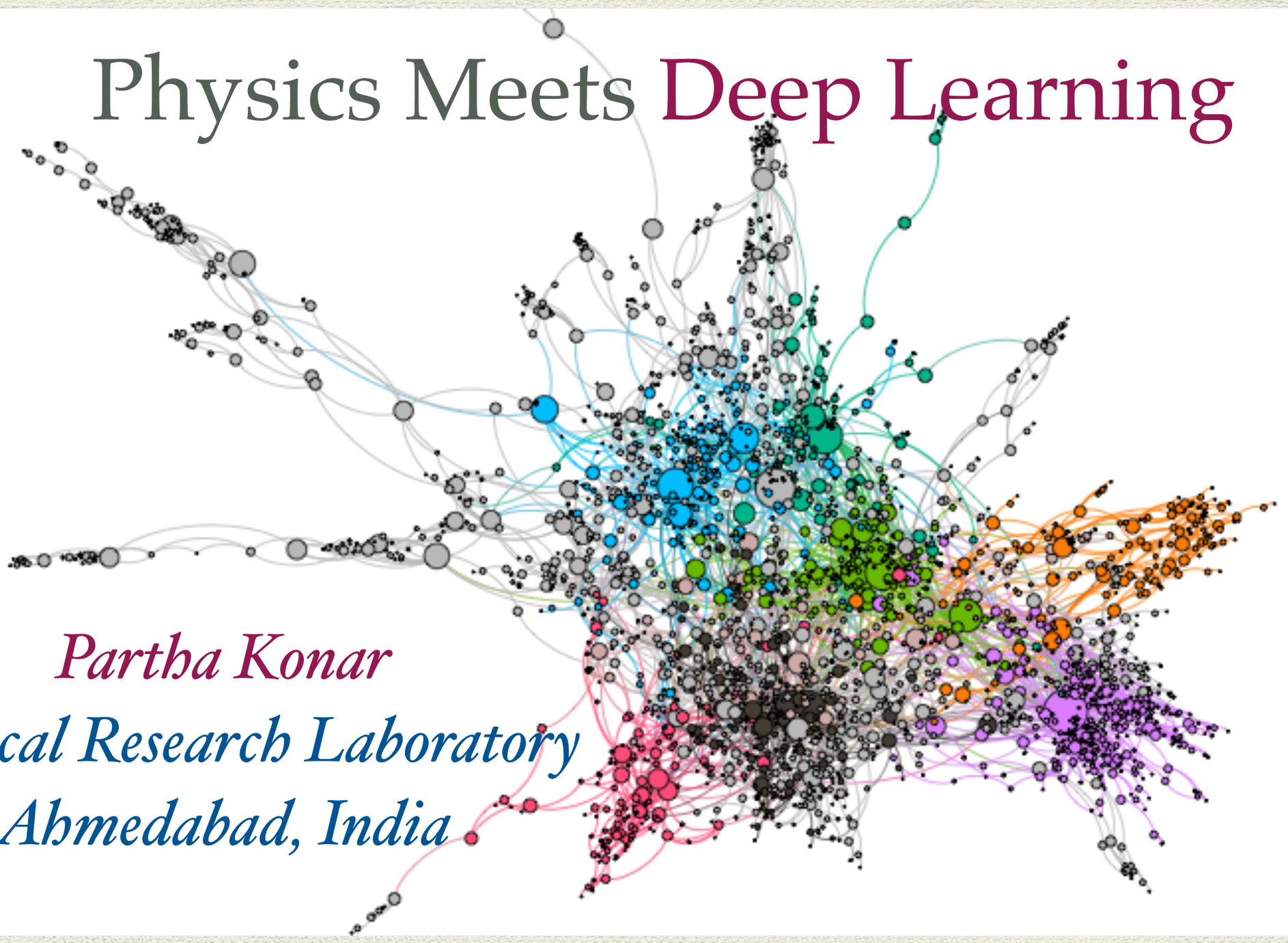
# Physics Meets Deep Learning




*Partha Konar*

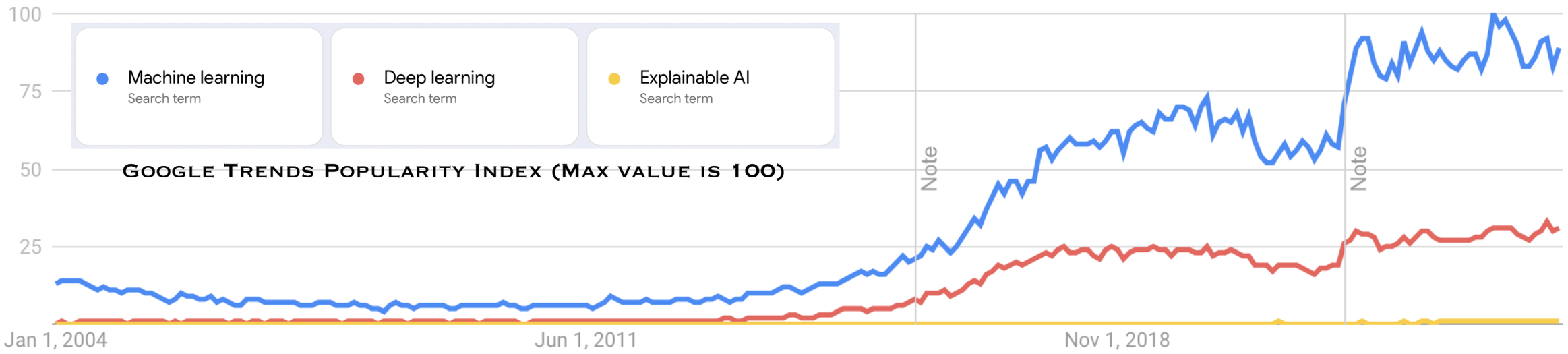
*Physical Research Laboratory*

*Ahmedabad, India*

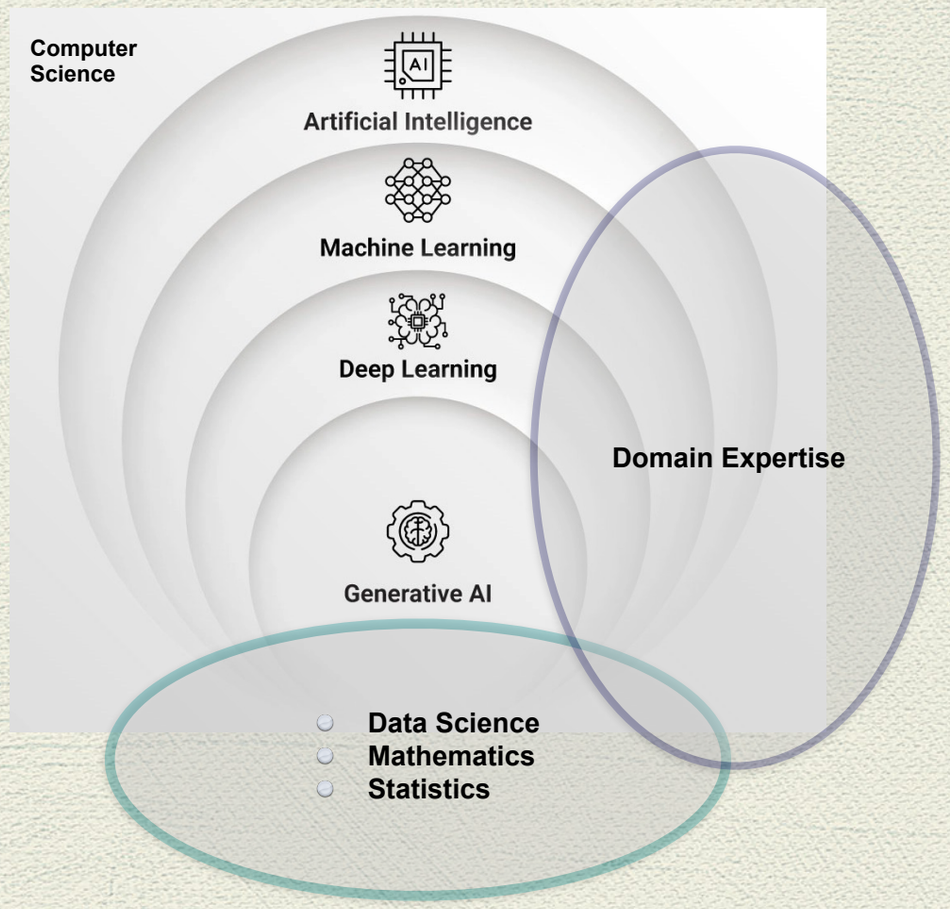

SRM UNIVERSITY AP
Andhra Pradesh

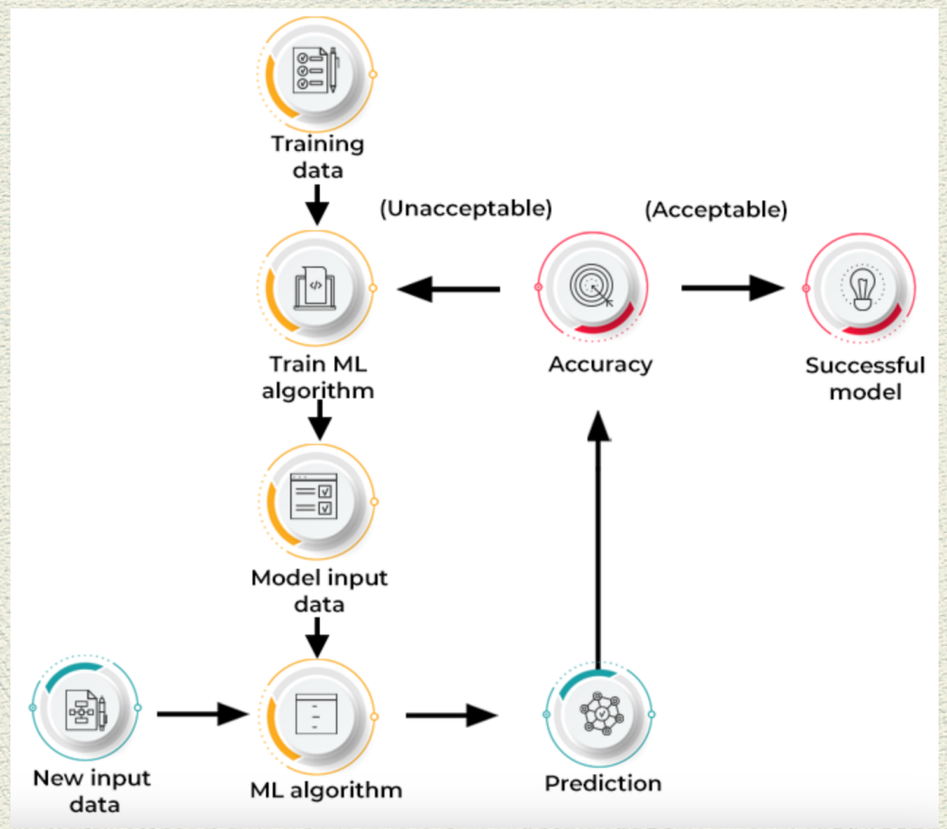
FROM BIG BANG TO NOW :
A THEORY-EXPERIMENT DIALOGUE

January 23 - 25, 2025

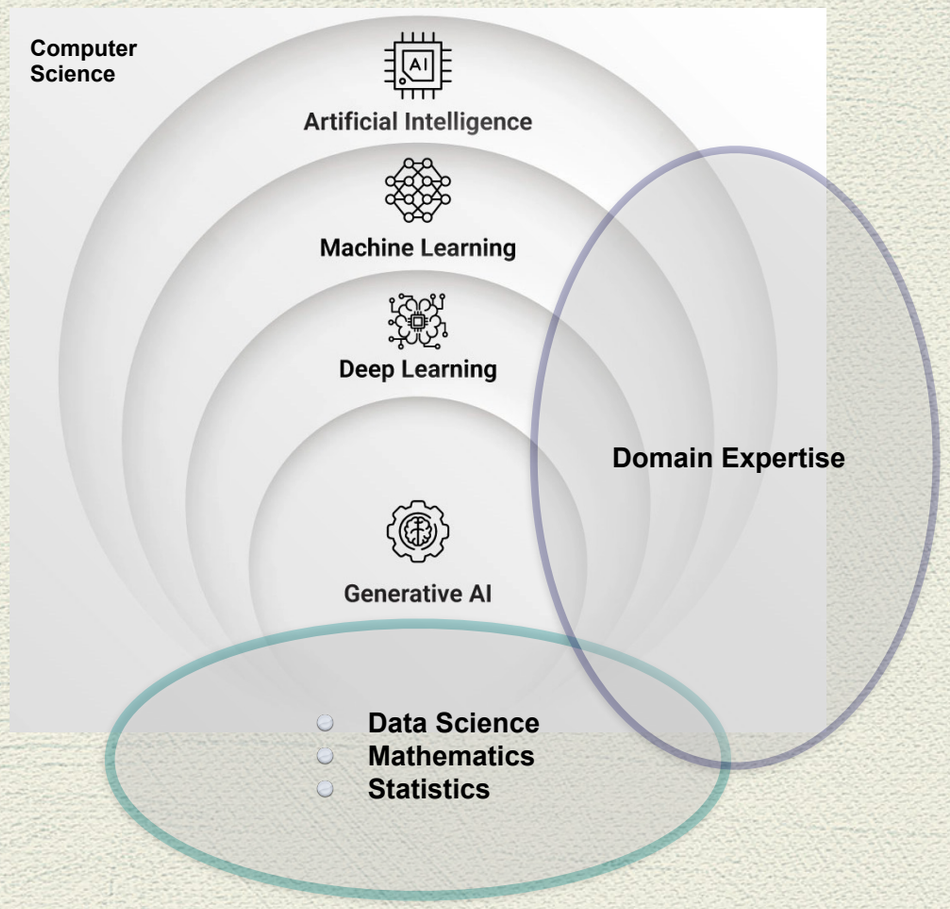

Machine learning
Search term
Deep learning
Search term
Explainable AI
Search term
Google Trends Popularity Index (Max value is 100)
100
75
50
25
Note
Note
Jan 1, 2004
Jun 1, 2011
Nov 1, 2018

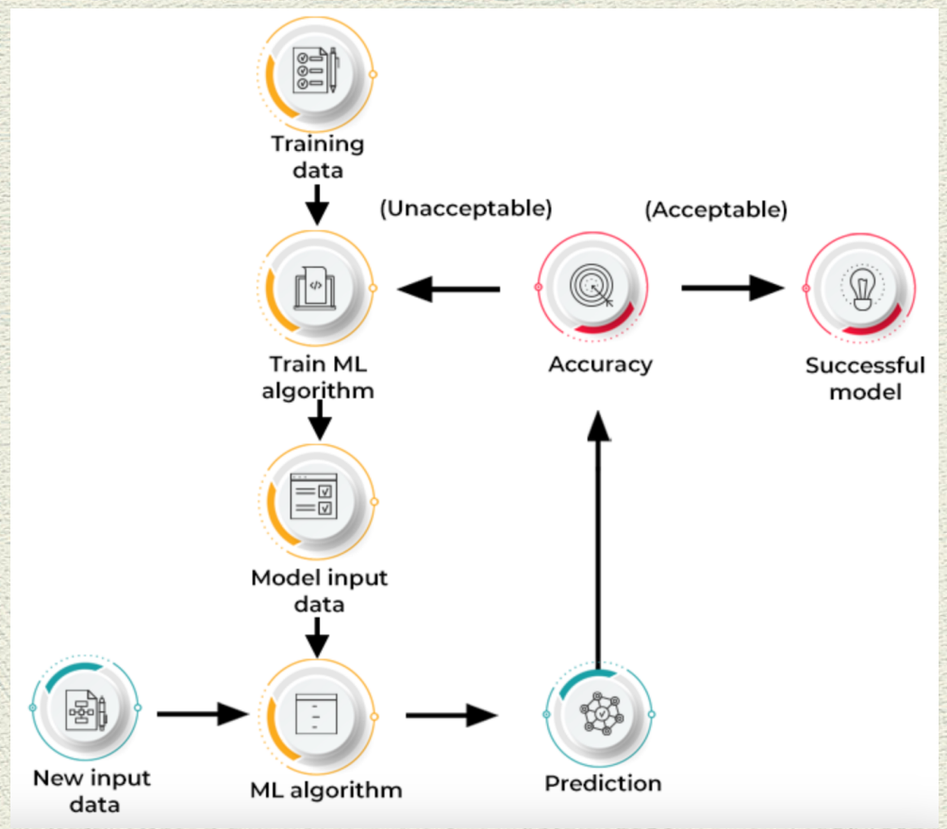
# How ML works?


Computer Science
Artificial Intelligence
Machine Learning
Deep Learning
Generative AI
Domain Expertise
Data Science
Mathematics
Statistics


Training data
(Unacceptable)
(Acceptable)
Train ML algorithm
Accuracy
Successful model
Model input data
New input data
ML algorithm
Prediction

# How ML works?




- **Decision boundary:** surface that divides multi-dim feature space into distinct groups of data points.
- **Training:** ML algorithm discovers the decision boundary
- **Testing**: Then uses to forecast the class of unseen data points.
- Key drivers for its growth

  — **Data, Algorithms & Hardware (graphics processing unit or GPU)**

# Three Ways to Learn

## Supervised

## Unsupervised

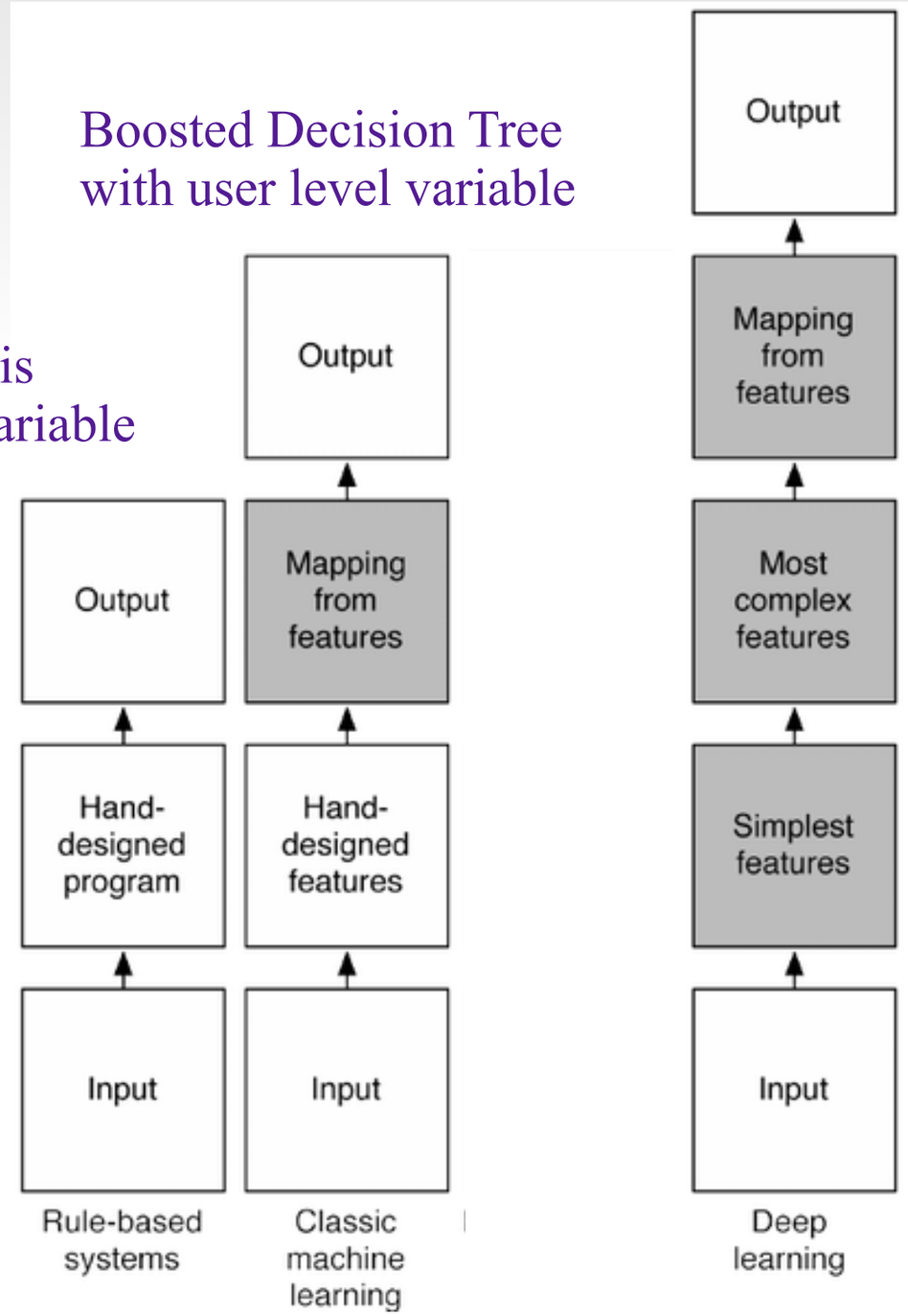
## Reinforcement

# Work Flow

- Deep Machine Learning : **Power and flexibility** by representing the complex world as a **nested hierarchy of concepts** -> Each concept defined in relation to simpler concepts
- DNN outperformed other machine learning and hand-designed functionality

Boosted Decision Tree
with user level variable

Cut based analysis
with user level variable



# MACHINE LEARNING

## AND .. GOING DEEPER

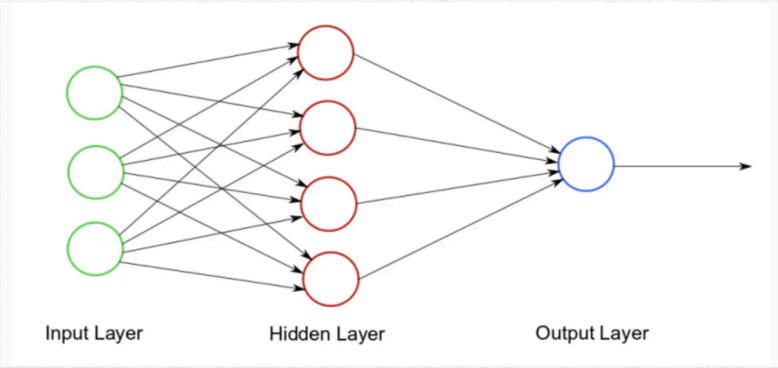
- Universal function approximation: NN with a single hidden layer can approximate any continuous function to any desired precision!

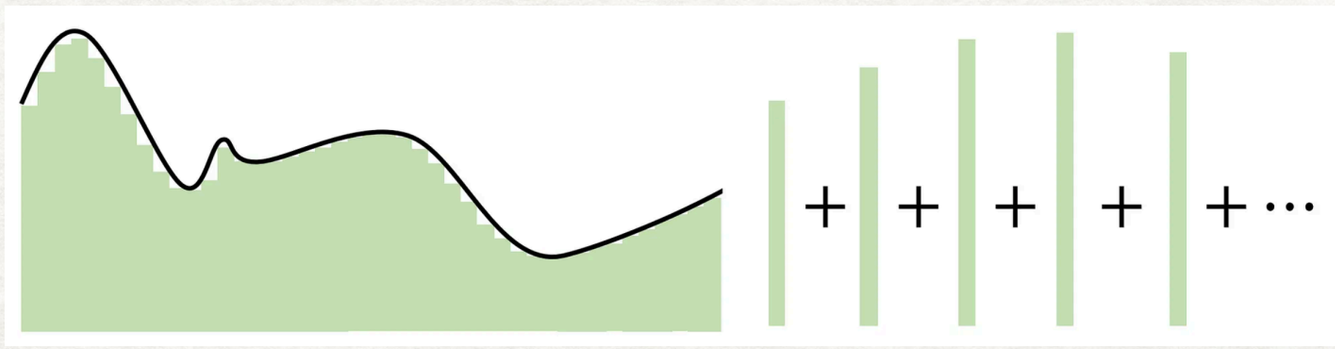
- Deep learning models with **multiple hidden layers** solves the need for infinitely large no of nodes in shallow NN

- Learning **scalable** with data - larger data for better performance

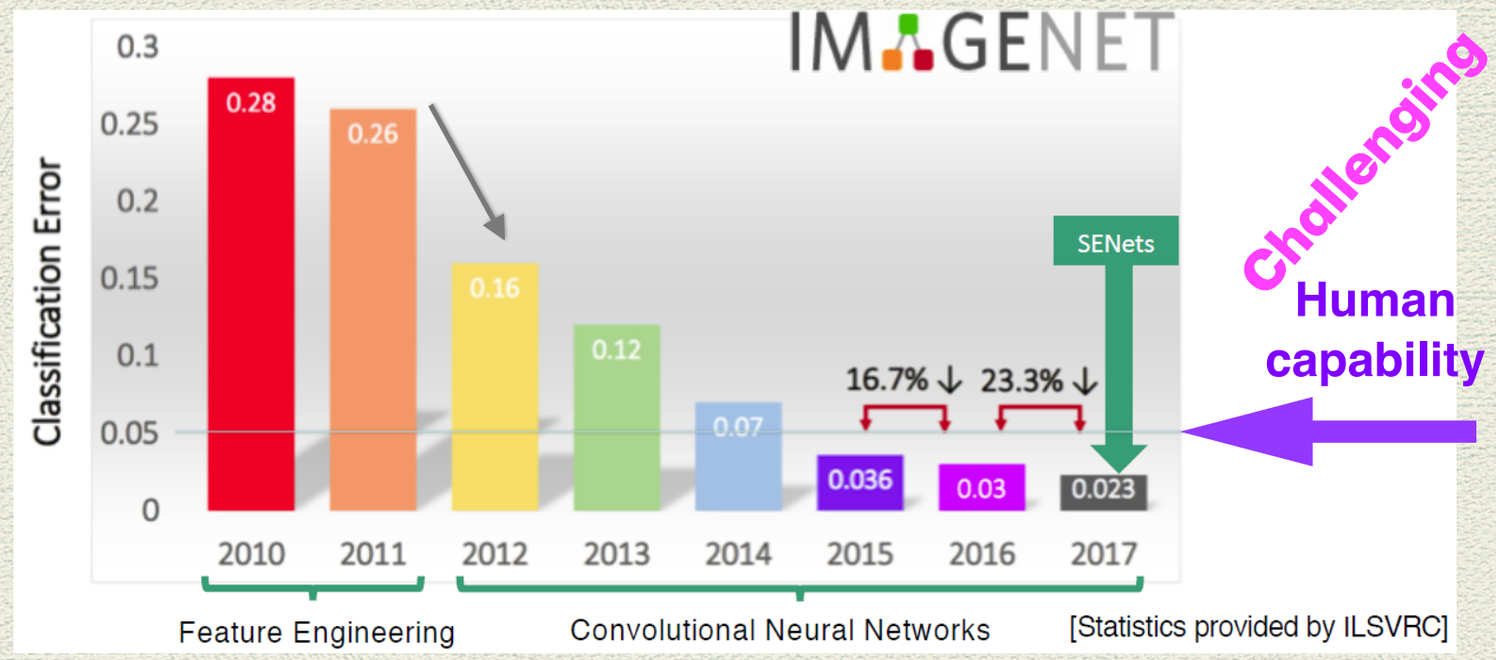
- Deep learning models are now capable of **extracting feature directly from low level data**

  — End for physics intuitive high level variables from domain experts?



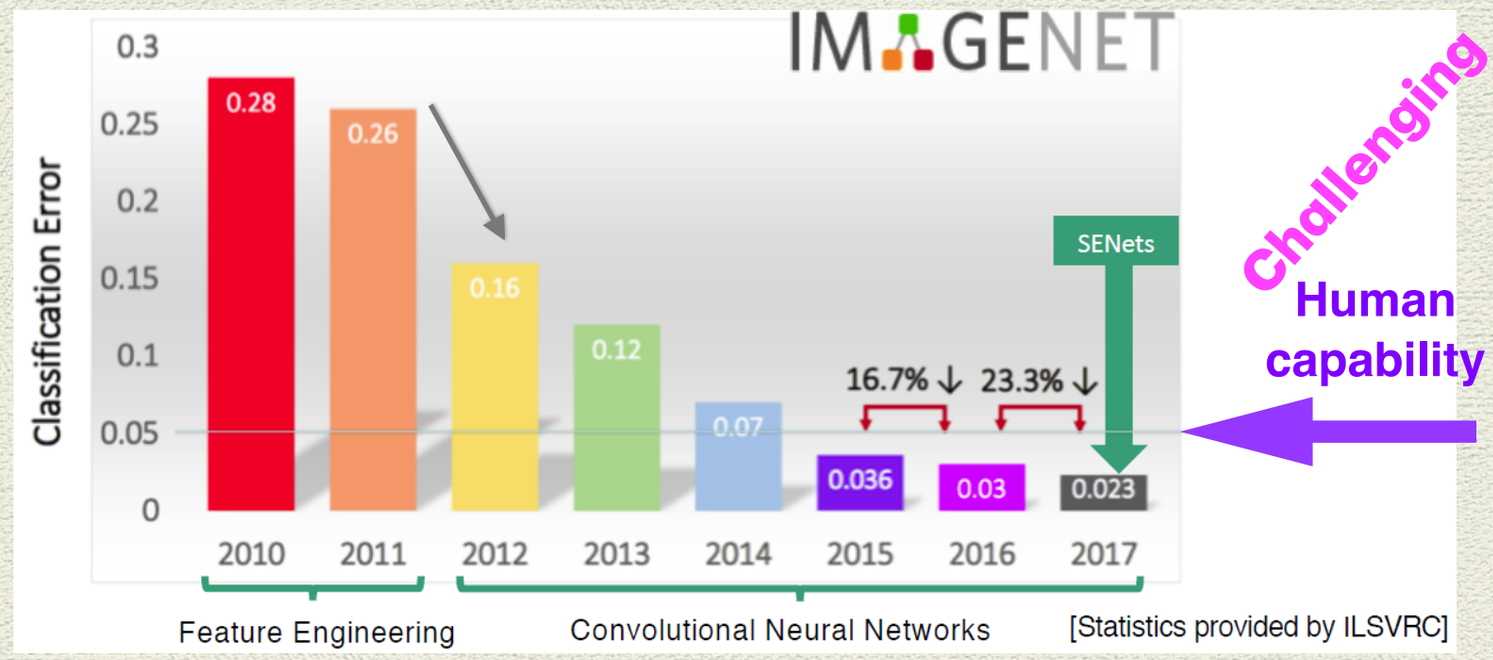
# Progress of Deep Learning





- ImageNet - Large Scale Visual Recognition Challenge (ILSVRC) held each year : largest contest in object recognition
- 2012 - AlexNet [Deep CNN by Alex Krizhevsky etal] ~ 15.4% error (2nd 26.2%!)]
- Since then, these competitions are consistently won by deep convolutional nets

# Progress of Deep Learning

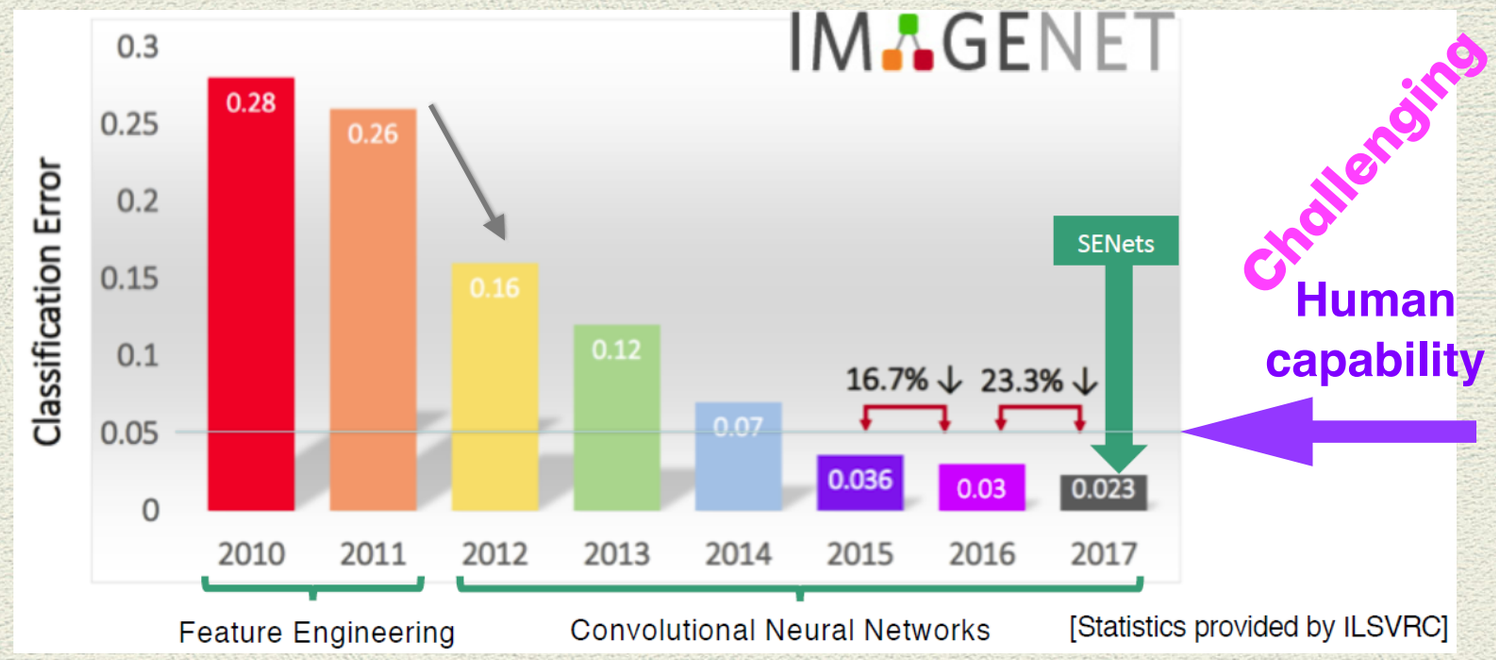




- ImageNet - Large Scale Visual Recognition Challenge (ILSVRC) held each year : largest contest in object recognition
- 2012 - AlexNet [Deep CNN by Alex Krizhevsky etal] ~ 15.4% error (2nd 26.2%!)]
- Since then, these competitions are consistently won by deep convolutional nets

With Ilya Sutskever &
PhD advisor Geoffrey Hinton

# Progress of Deep Learning



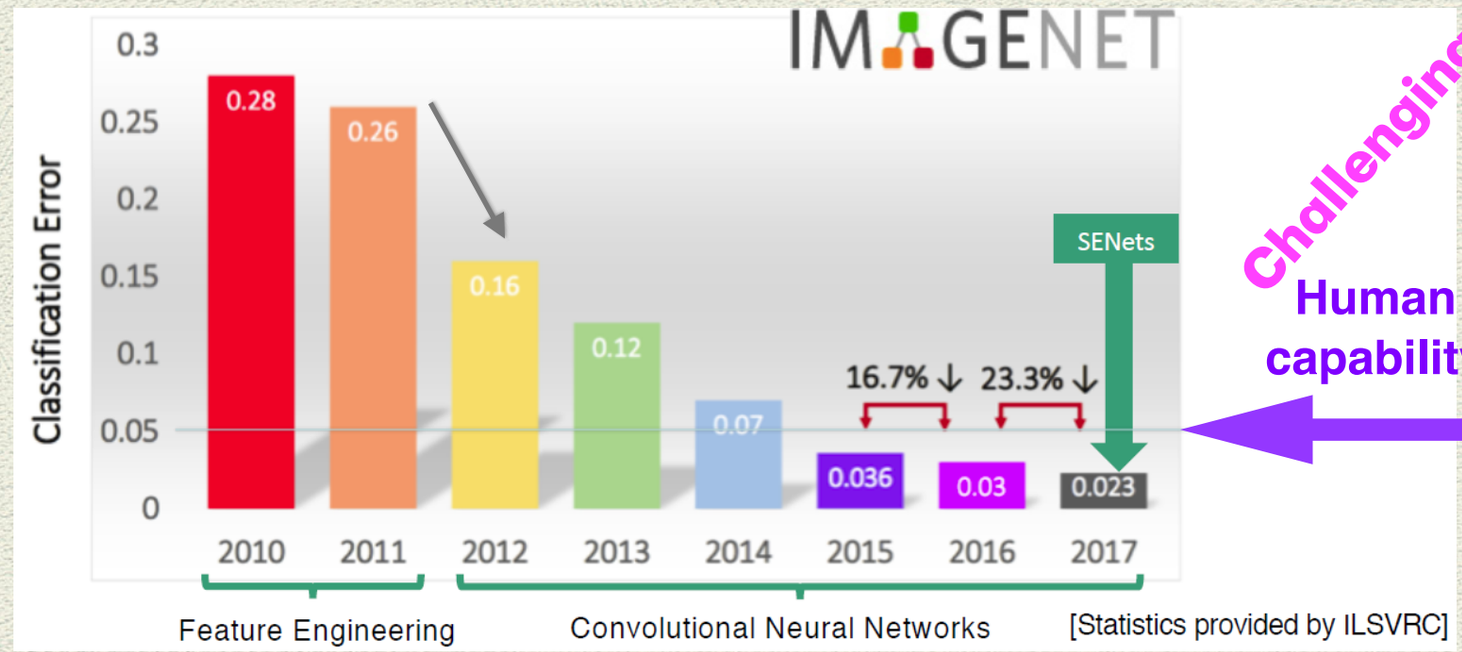


- ImageNet - Large Scale Visual Recognition Challenge (ILSVRC) held each year : largest contest in object recognition
- 2012 - AlexNet [Deep CNN by Alex Krizhevsky etal] ~ 15.4% error (2nd 26.2%!)]
- Since then, these competitions are consistently won by deep convolutional nets

With Ilya Sutskever & PhD advisor Geoffrey Hinton

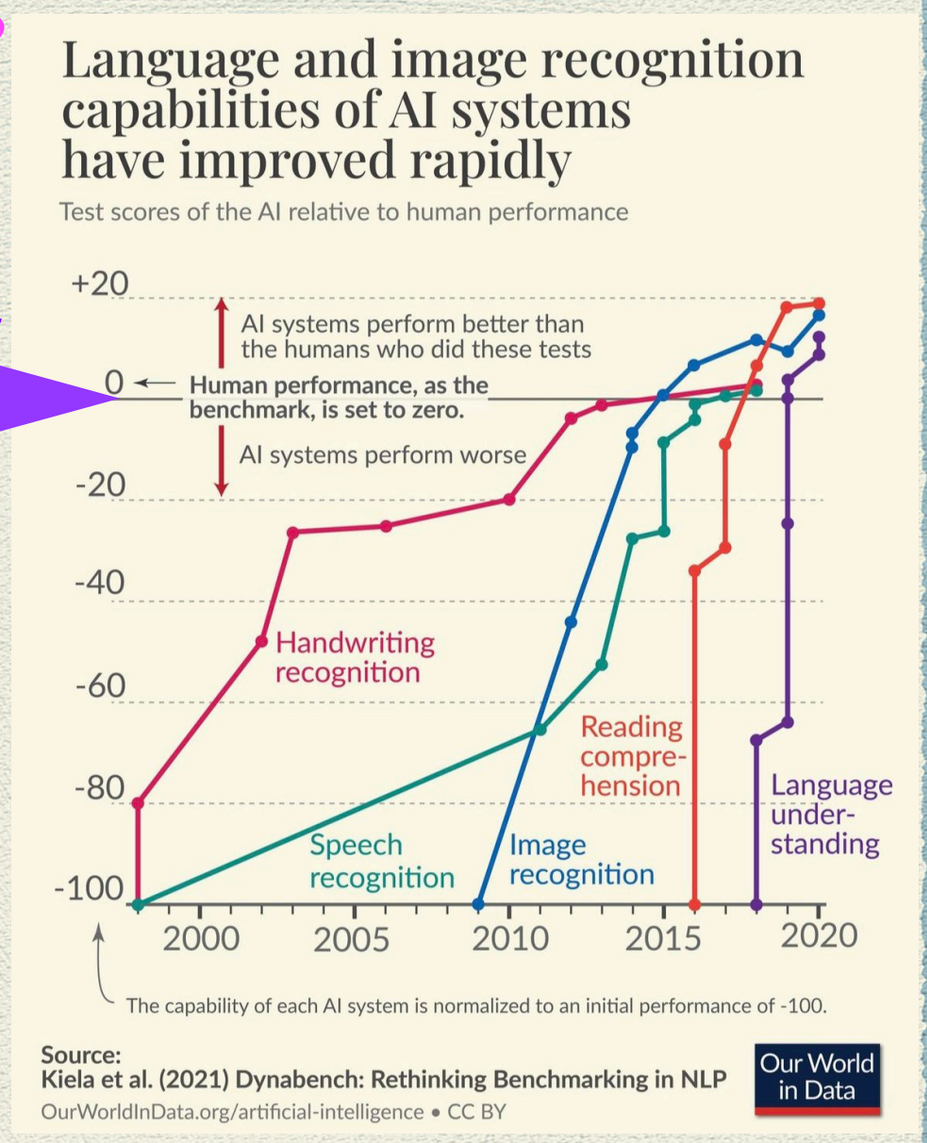
- Imitate (defeat) human intelligence and capability in **visual perception, speech recognition, decision-making, language processing, and so on.**

# Progress of Deep Learning




Challenging

Human capability

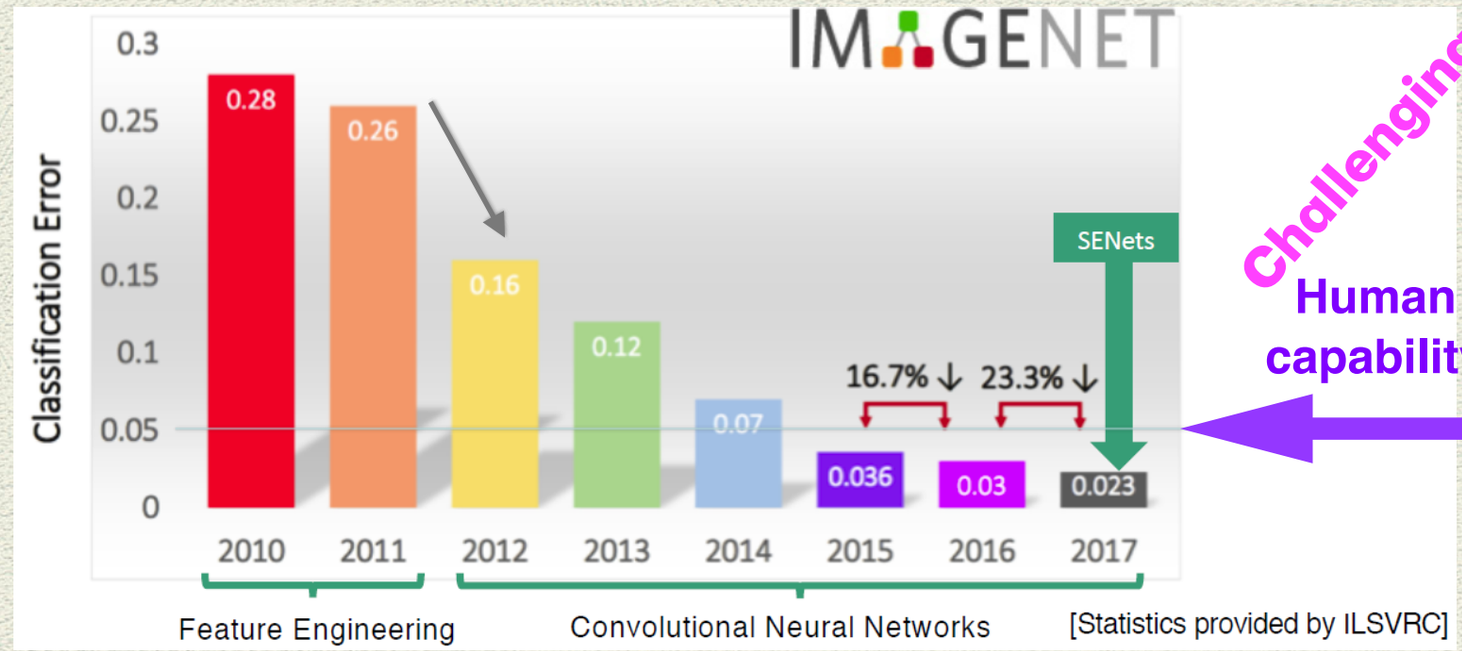



- ImageNet - Large Scale Visual Recognition Challenge (ILSVRC) held each year : largest contest in object recognition
- 2012 - AlexNet [Deep CNN by Alex Krizhevsky etal] ~ 15.4% error (2nd 26.2%!)]
- Since then, these competitions are consistently won by deep convolutional nets

- Imitate (defeat) human intelligence and capability in **visual perception, speech recognition, decision-making, language processing, and so on.**

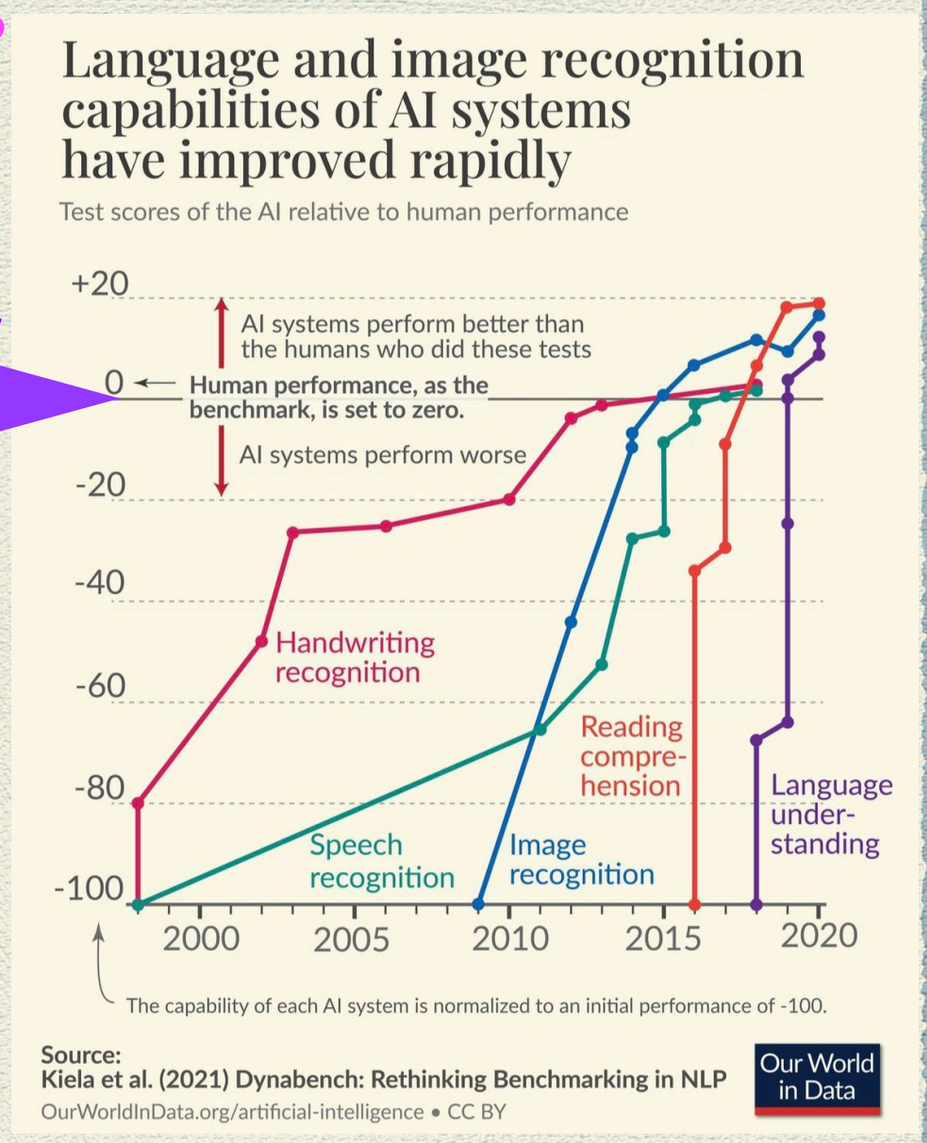
# Progress of Deep Learning

◉ **New era of Artificial General intelligence (AGI)**




Challenging

Human capability

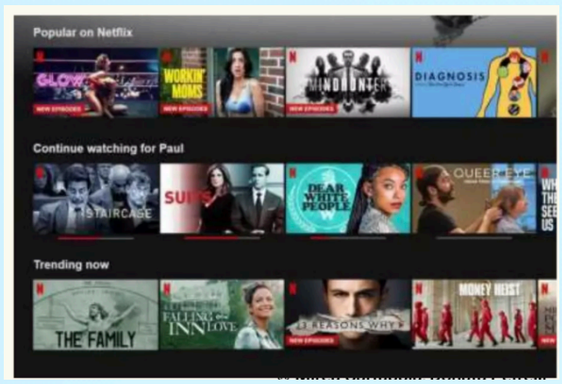



- ImageNet - Large Scale Visual Recognition Challenge (ILSVRC) held each year : largest contest in object recognition
- 2012 - AlexNet [Deep CNN by Alex Krizhevsky etal] ~ 15.4% error (2nd 26.2%!)]
- Since then, these competitions are consistently won by deep convolutional nets

◉ Imitate (defeat) human intelligence and capability in
**visual perception, speech recognition, decision-making, language processing, and so on.**

*While world is mesmerised with different impossibles*
*done with ML applications in our everyday life,*

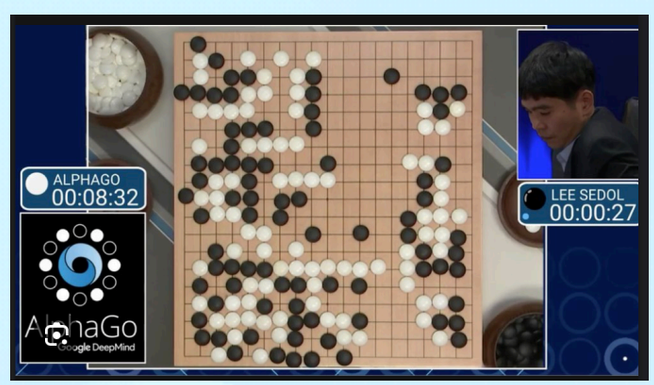
## Content Recommendation



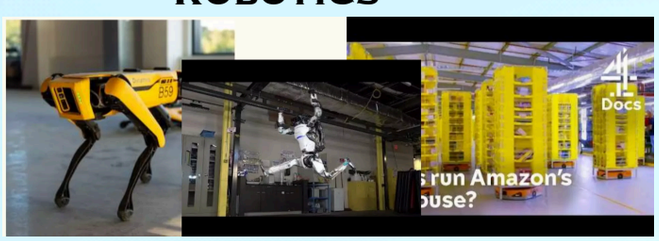

## Self-driving Car

## Face Recognition

## Gaming




## Robotics




## Creativity Culture music

## Virtual Assistance


**Artificial Intelligence in Everyday Apps**

Predictive Search

Object Detection

News Feed Relevance

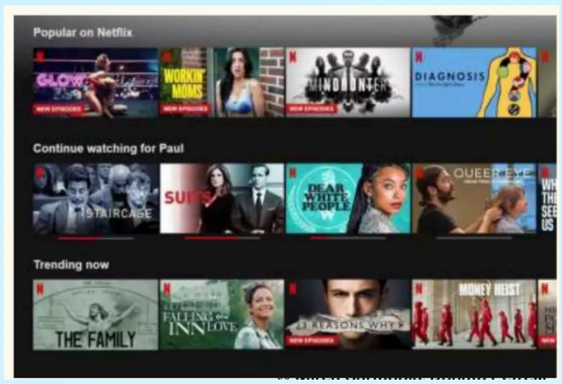
Recommendations

Matching Algorithm

Smart Replies

*While world is mesmerised with different impossibles*
*done with ML applications in our everyday life,*

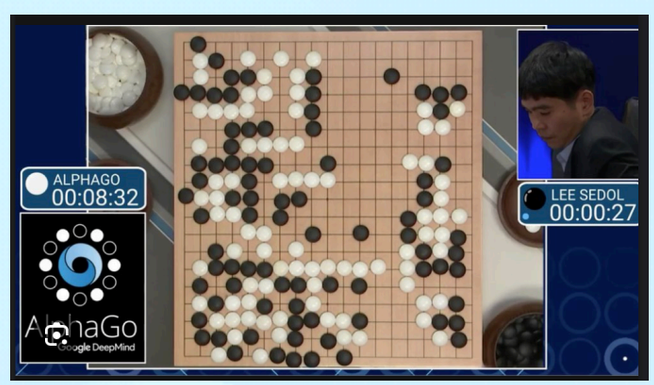
### Content Recommendation



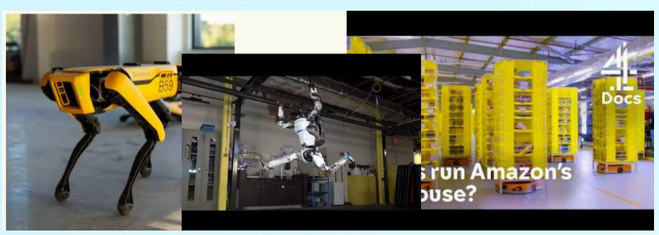

### Self-driving Car

### Face Recognition

### Gaming




### Robotics



### Creativity Culture music

### Virtual Assistance

ChatGPT Gemini

**Artificial Intelligence in Everyday Apps**

Predictive Search

Object Detection

News Feed Relevance

Recommendations

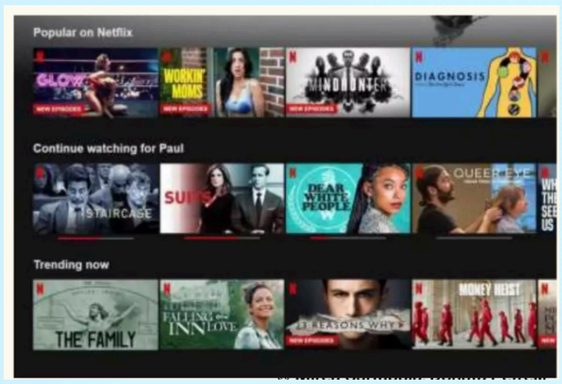
Matching Algorithm

Smart Replies

*Dramatic shifts are also happening in almost all research fields —*
*including Healthcare, Medicine, Finance, Education services etc*

*While world is mesmerised with different impossibles*
*done with ML applications in our everyday life,*

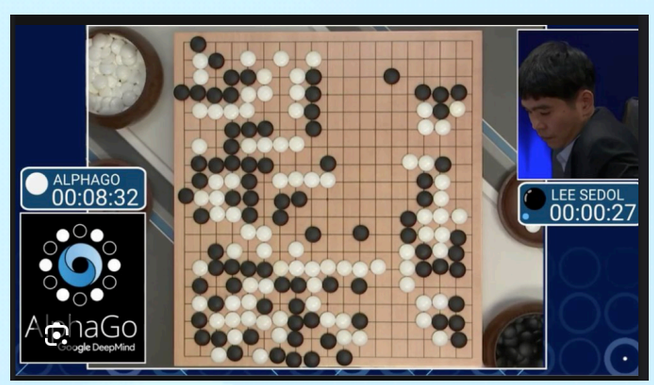
### Content Recommendation



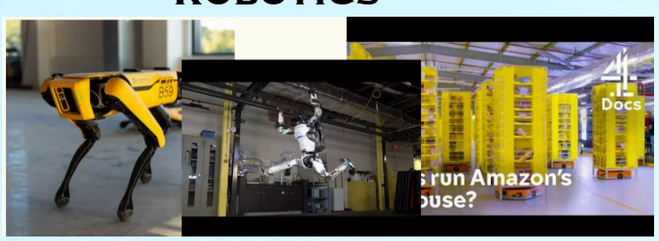

### Self-driving Car

### Face Recognition

### Gaming




### Robotics




ChatGPT Gemini

**Artificial Intelligence in Everyday Apps**

### Creativity Culture music

### Virtual Assistance

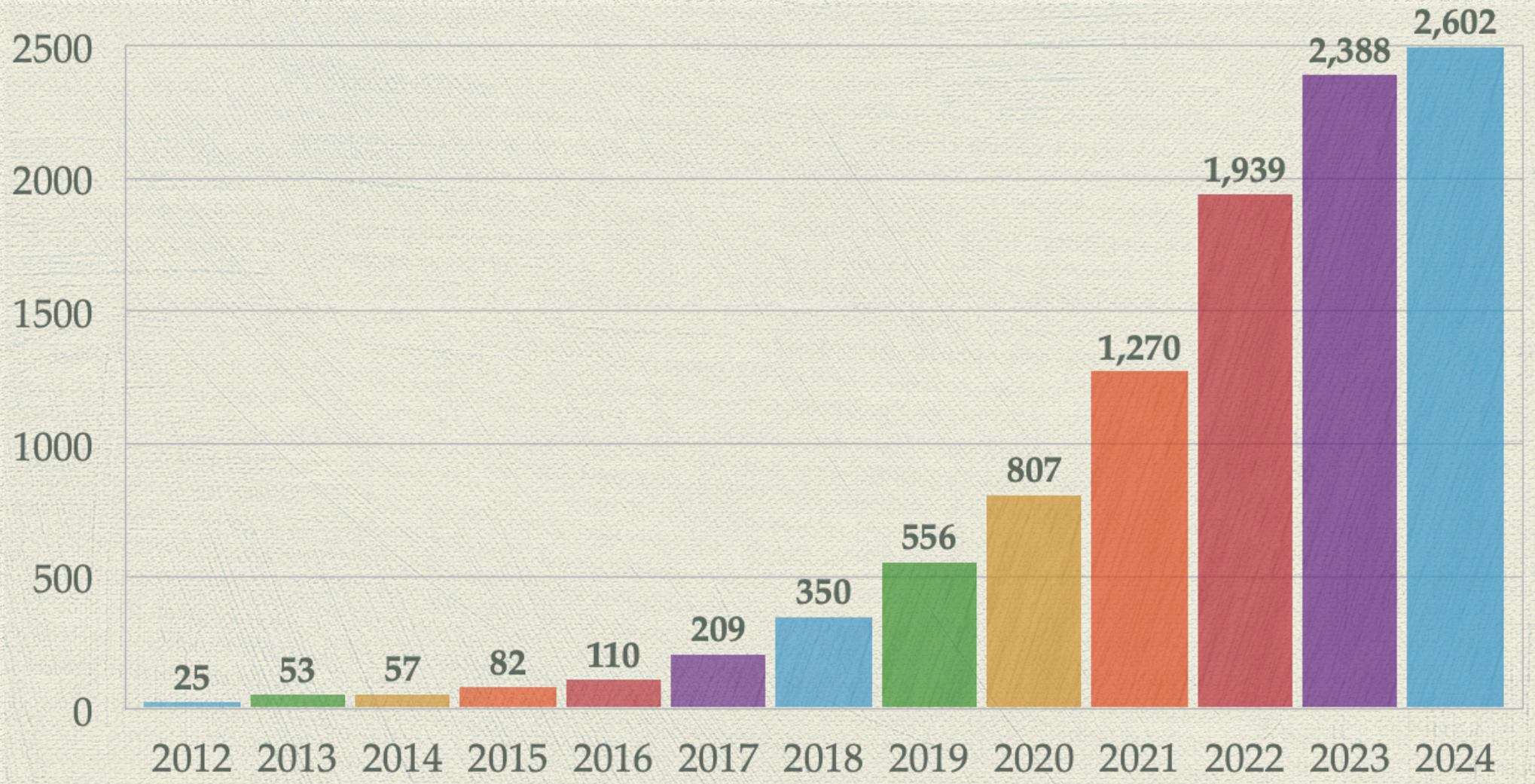
Predictive Search | Object Detection | News Feed Relevance

Recommendations | Matching Algorithm | Smart Replies

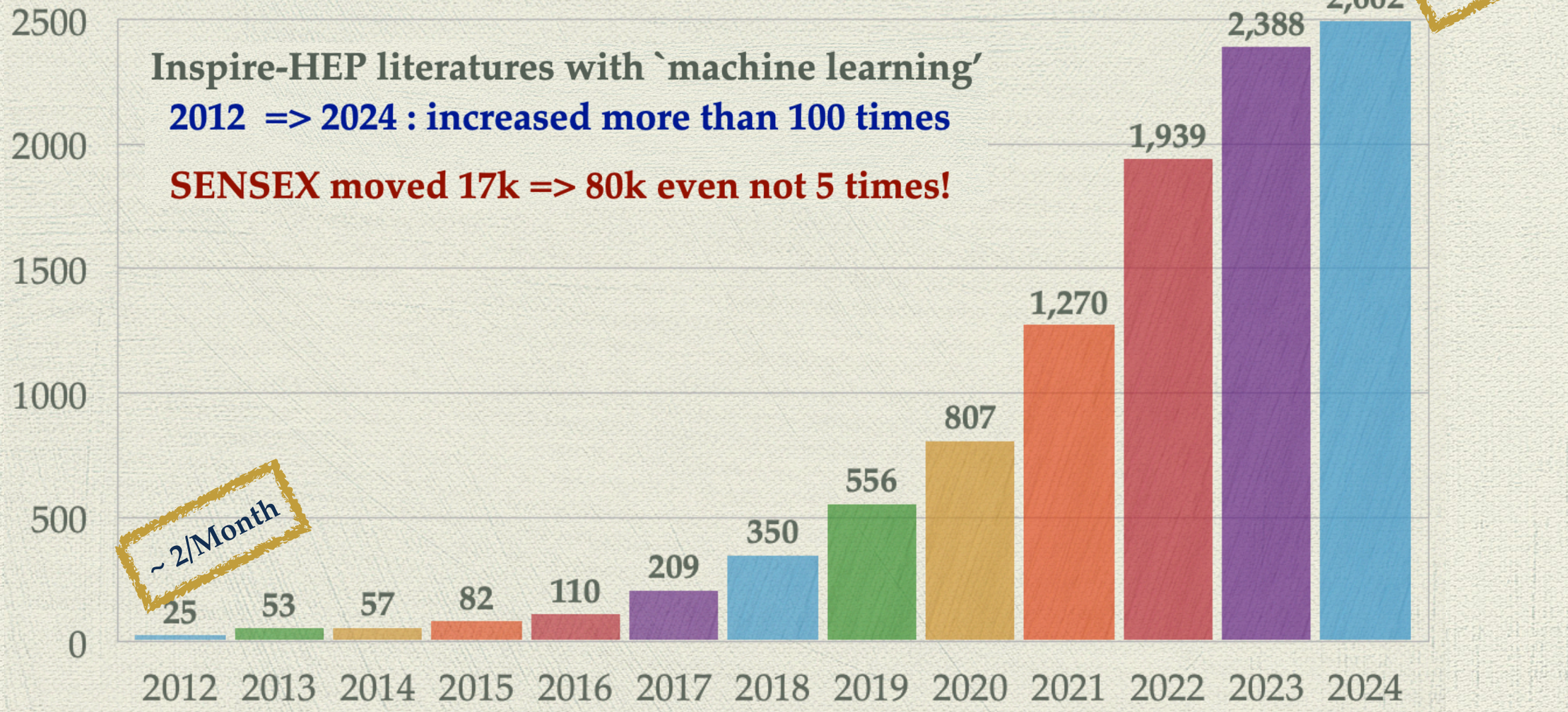
*Dramatic shifts are also happening in almost all research fields —*
*including Healthcare, Medicine, Finance, Education services etc*

*Several experimental results found their relevance — such scientific discoveries are ML driven*


Google Trends Popularity Index (Max value is 100)




Google Trends Popularity Index (Max value is 100)



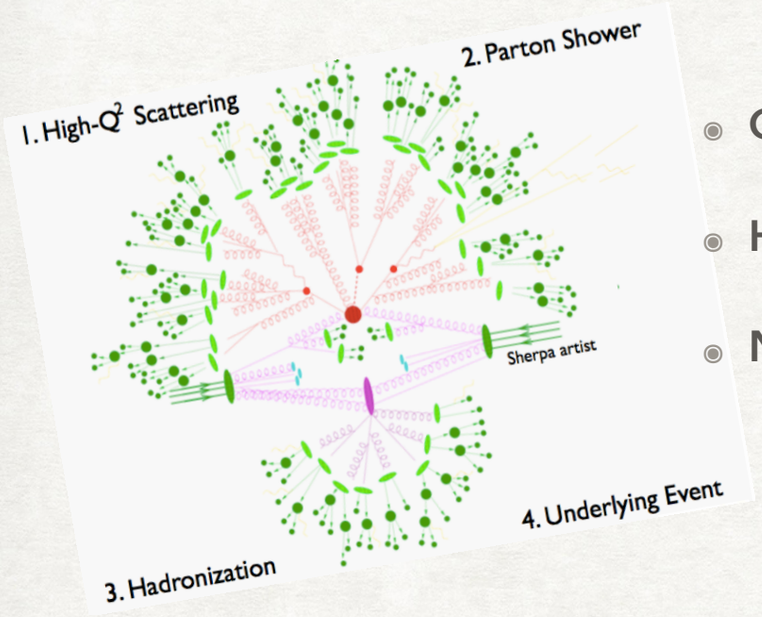
# DEEP MACHINE LEARNING

## CATEGORY

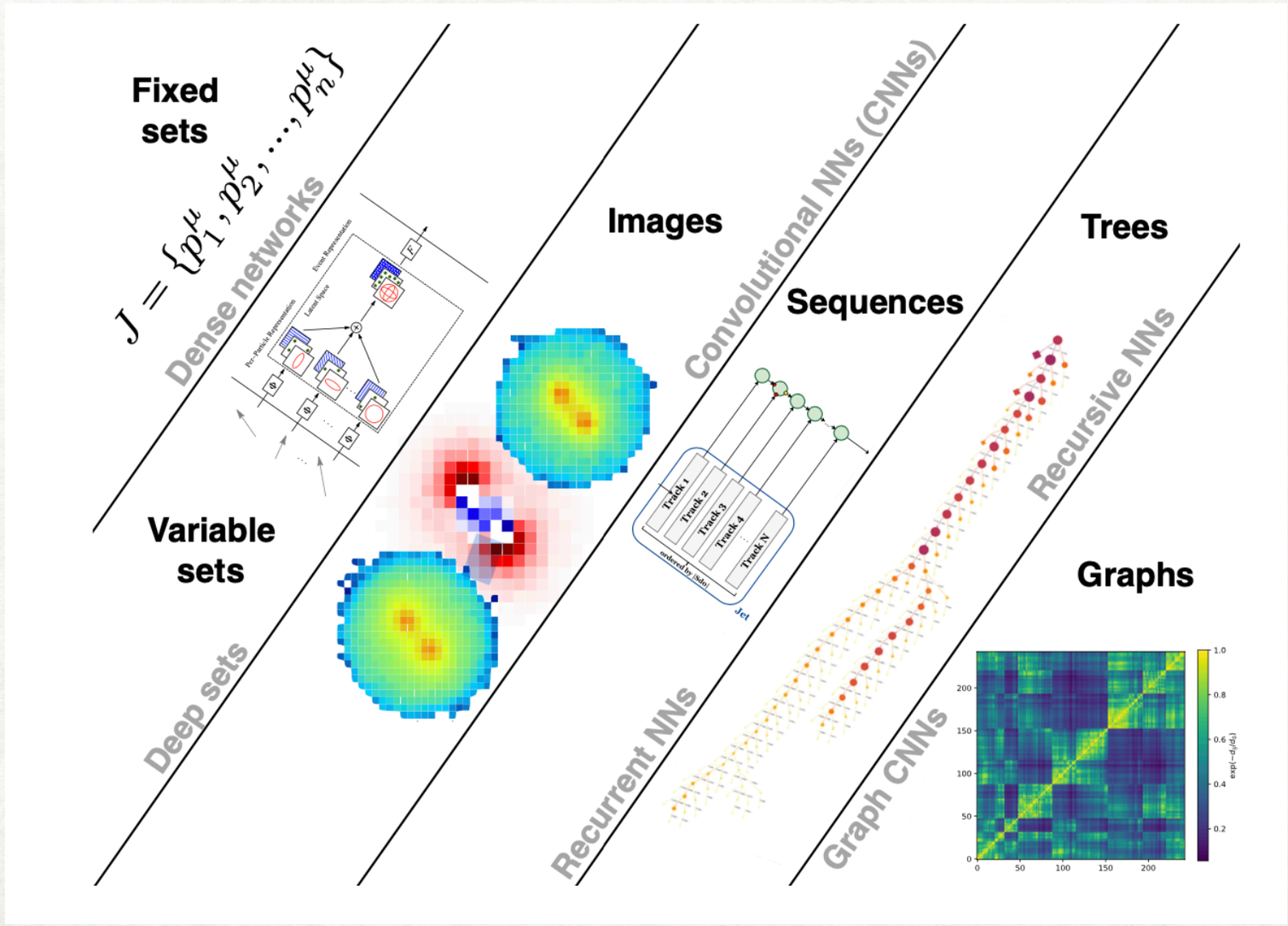
| Strategy | Representations | Targets / tagging | strategies |
|---|---|---|---|
| Classification | ➡ Jet Image<br>➡ Event Image<br>➡ Sequence (Recurrent NN)<br>➡ Graph (Graph NN)<br>➡ Sets (Point cloud - Graph) | ➡ Quarks vs gluons<br>➡ Boosted H / W / Z / Top tag<br>➡ New particles and models<br>➡ Particle tagging at detector<br>➡ Neutrino flavour | ➡ Weak/ Semi/ Un-supervised<br>➡ Reinforcement Learning<br>➡ Quantum Machine Learn<br>➡ Feature Ranking<br>➡ Optimal Transport |
| Regression | ➡ Parameter estimation<br>➡ Pileup mitigation<br>➡ Parton Distribution Func<br>➡ Symbolic Regression<br>➡ Function Approximation | | |
| Generative models | | ➡ GANs<br>➡ Autoencoders<br>➡ Phase space generation<br>➡ Normalizing flows | |
| Anomaly detection | | | |

# JET REPRESENTATION

## JET DATA - IMAGES, SEQUENCES AND SETS

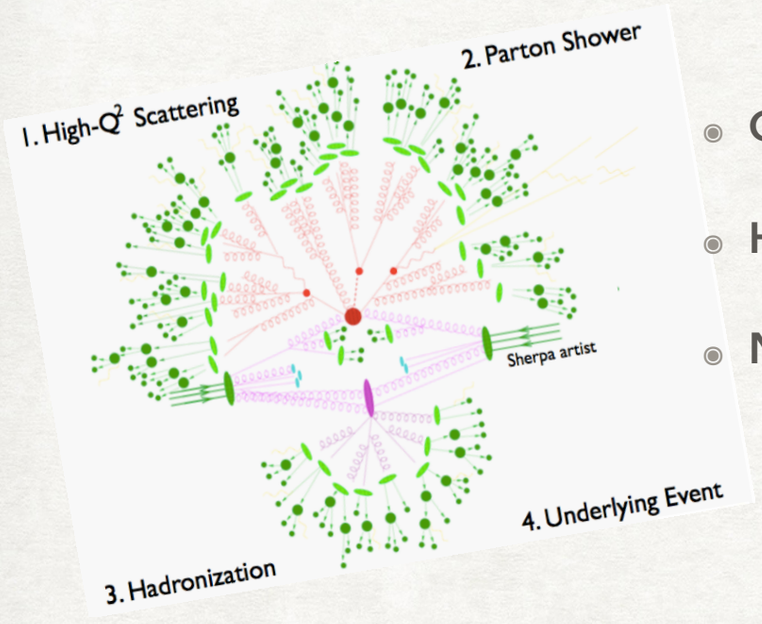



- QCD Jets have a rich & complex structure - perfect playing field
- How related to the first principles in Quantum Chromodynamics?
- No unique way for encoding radiation pattern into a particular data structure



# JET REPRESENTATION

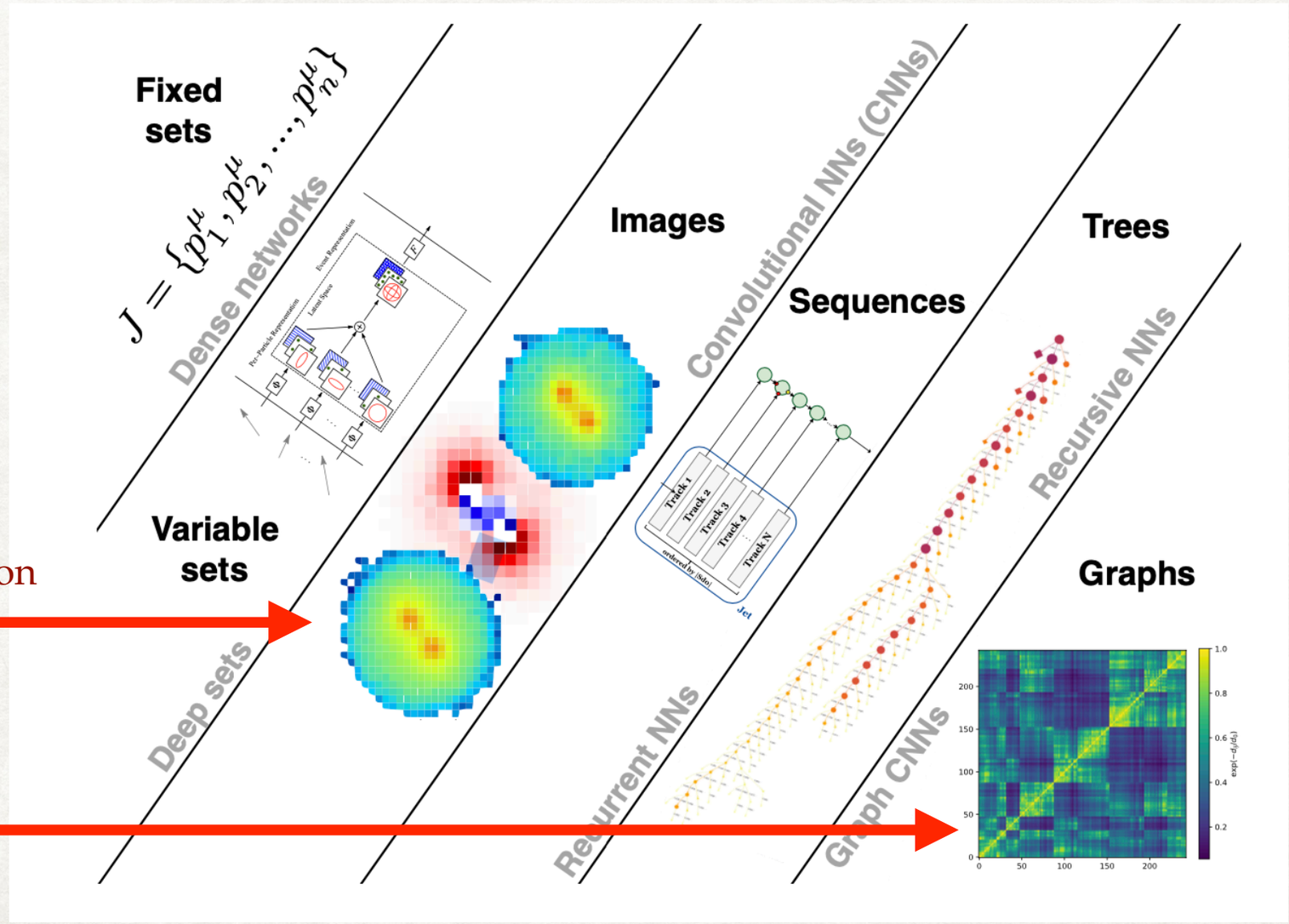
## JET DATA - IMAGES, SEQUENCES AND SETS



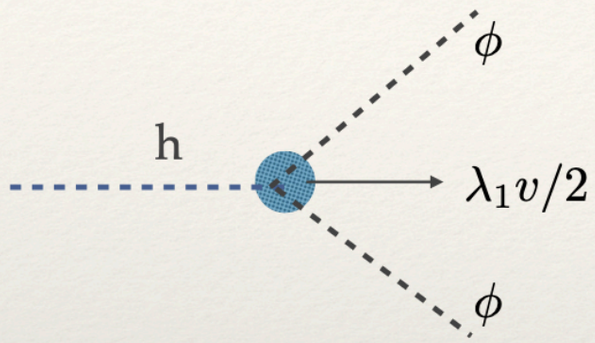

- QCD Jets have a rich & complex structure - perfect playing field
- How related to the first principles in Quantum Chromodynamics?
- No unique way for encoding radiation pattern into a particular data structure

- Theoretically motivated Qn

- Low level image of jet and QCD radiation
  Using CNN network
- Hadronic jet and QCD radiation
  Using GNN network



# HIGGS INVISIBLE DECAY

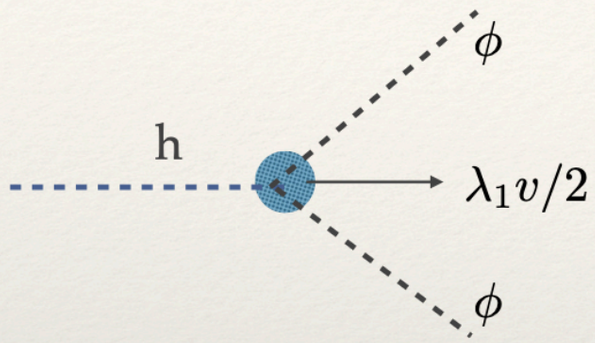



For $m_h < m_\phi/2$, Higgs can decay to a pair of Dark Matter.

**Decay width**

$$\Gamma_{h\to\phi\phi} = \frac{\lambda_1^2 v^2}{32\pi m_h^2}(m_h^2 - 4m_\phi^2)^{1/2}$$

# HIGGS INVISIBLE DECAY



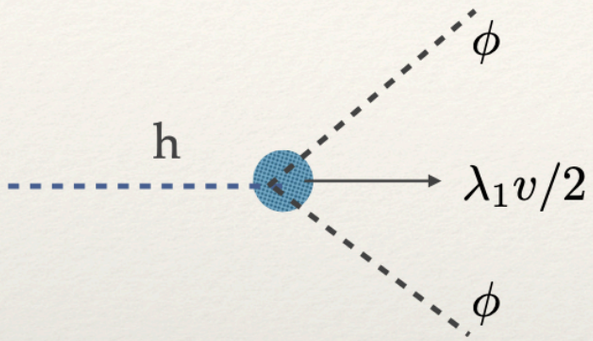

For $m_h < m_\phi/2$, Higgs can decay to a pair of Dark Matter.

Decay width

$$\Gamma_{h\to\phi\phi} = \frac{\lambda_1^2 v^2}{32\pi m_h^2}(m_h^2 - 4m_\phi^2)^{1/2}$$

**Strong limit => (Higgs portal) DM models constraint**

# HIGGS INVISIBLE DECAY

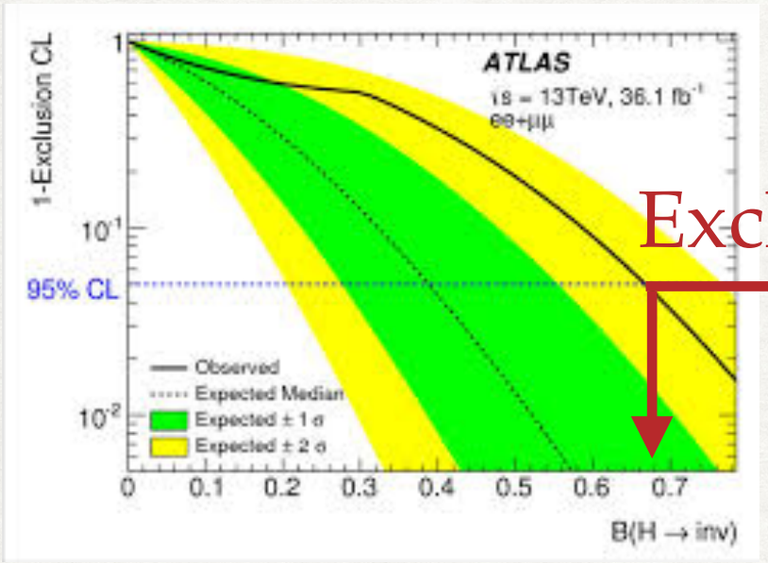



For $m_h < m_\phi/2$, Higgs can decay to a pair of Dark Matter.

Decay width

$$\Gamma_{h\to\phi\phi} = \frac{\lambda_1^2 v^2}{32\pi m_h^2}(m_h^2 - 4m_\phi^2)^{1/2}$$

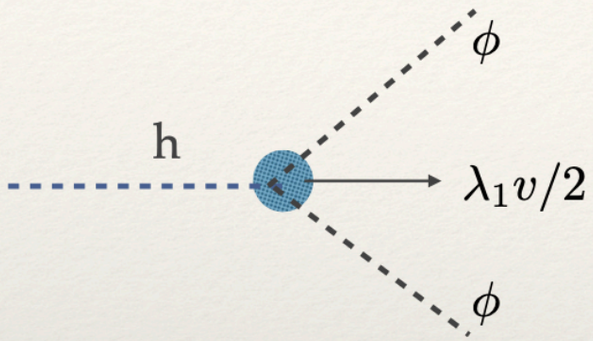
**Strong limit => (Higgs portal) DM models constraint**




Excluded

**Measured Br(h->inv) >> SM prediction (< 0.1%)!!**
**[23% (36fb), 10% (140fb)**

# HIGGS INVISIBLE DECAY

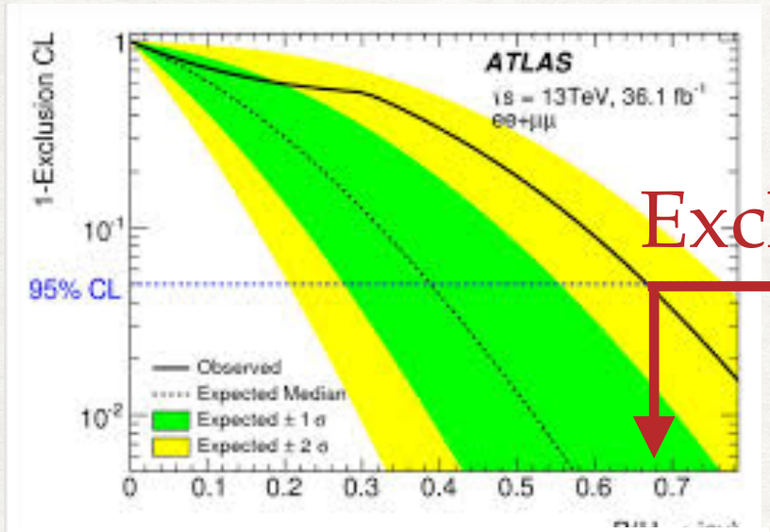



For $m_h < m_\phi/2$, Higgs can decay to a pair of Dark Matter.

Decay width

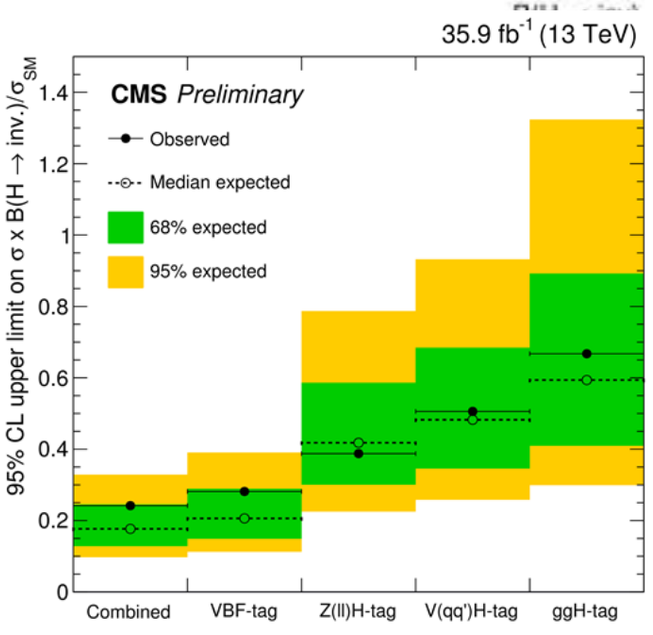
$$\Gamma_{h\to\phi\phi} = \frac{\lambda_1^2 v^2}{32\pi m_h^2}(m_h^2 - 4m_\phi^2)^{1/2}$$

**Strong limit => (Higgs portal) DM models constraint**

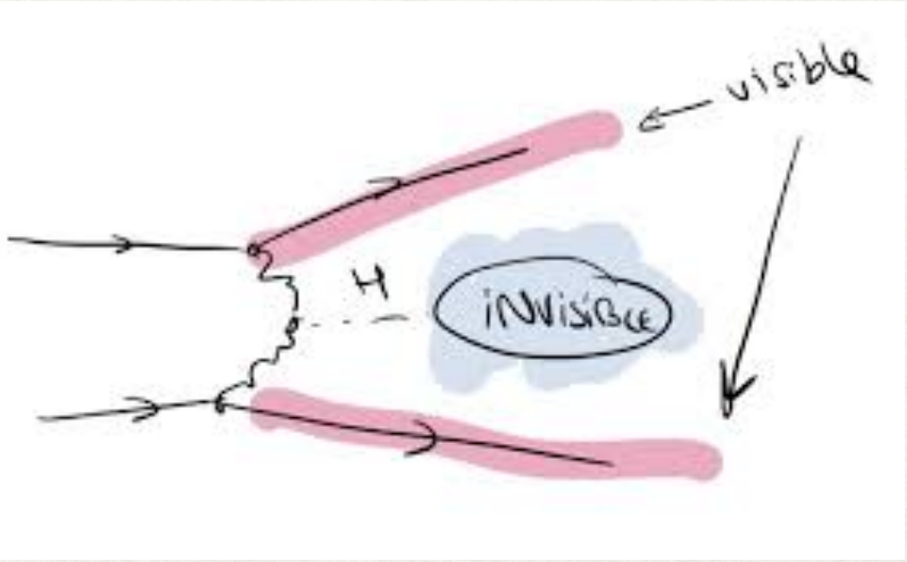



Excluded

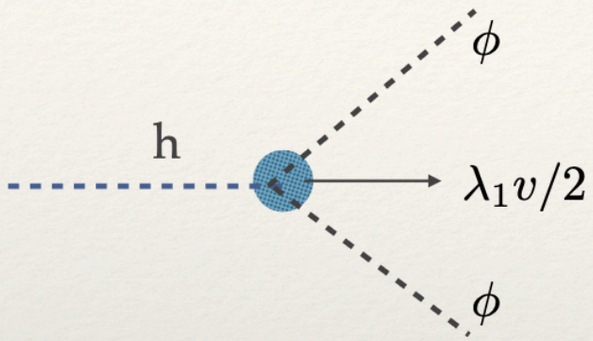
**Measured Br(h->inv) >> SM prediction (< 0.1%)!!**
**[23%(36fb), 10% (140fb)**




**VBF is most sensitive channel**



# HIGGS INVISIBLE DECAY

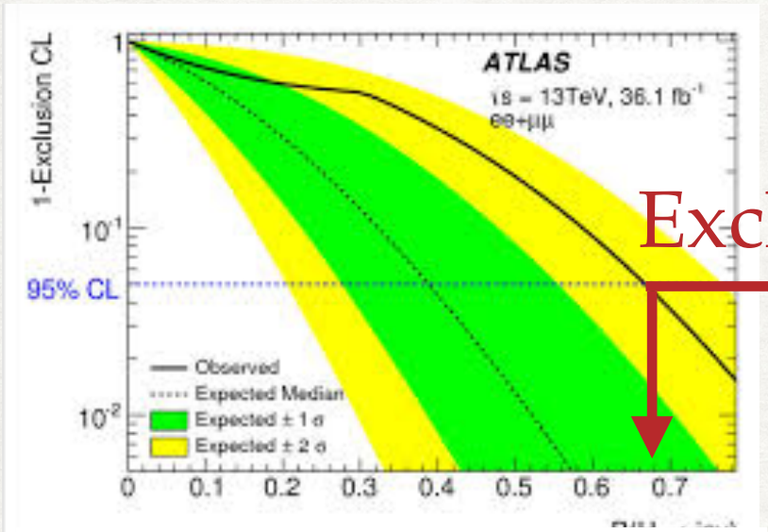



For $m_h < m_\phi/2$, Higgs can decay to a pair of Dark Matter.

Decay width

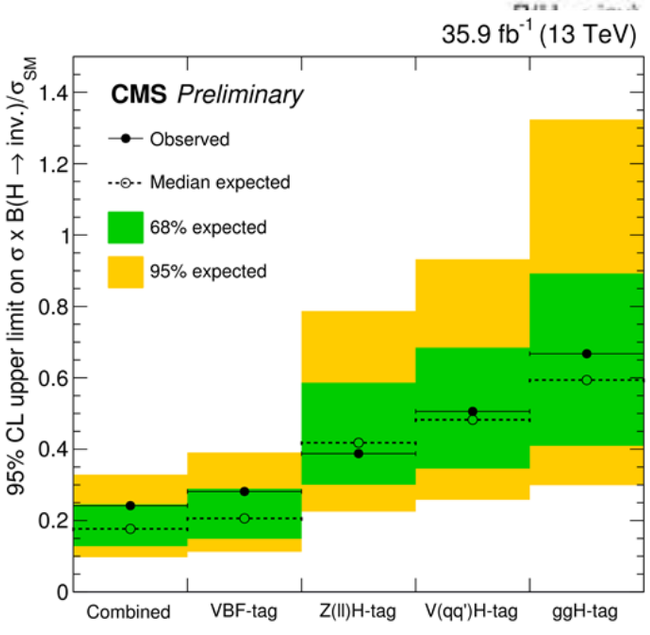
$$\Gamma_{h\to\phi\phi} = \frac{\lambda_1^2 v^2}{32\pi m_h^2}(m_h^2 - 4m_\phi^2)^{1/2}$$

**Strong limit => (Higgs portal) DM models constraint**

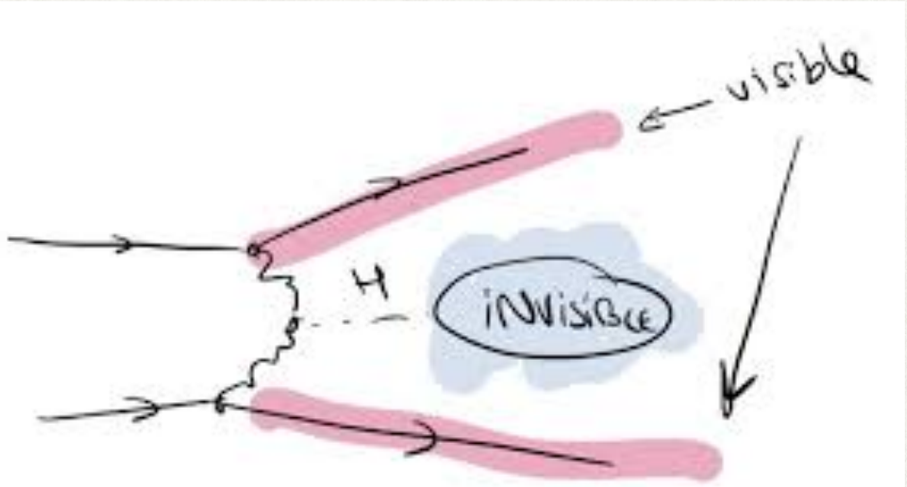



Excluded

**Measured Br(h->inv) >> SM prediction (< 0.1%)!!**
**[23%(36fb), 10% (140fb)**




**VBF is most sensitive channel**




**Can CNN learn characteristics of VBF signal?**

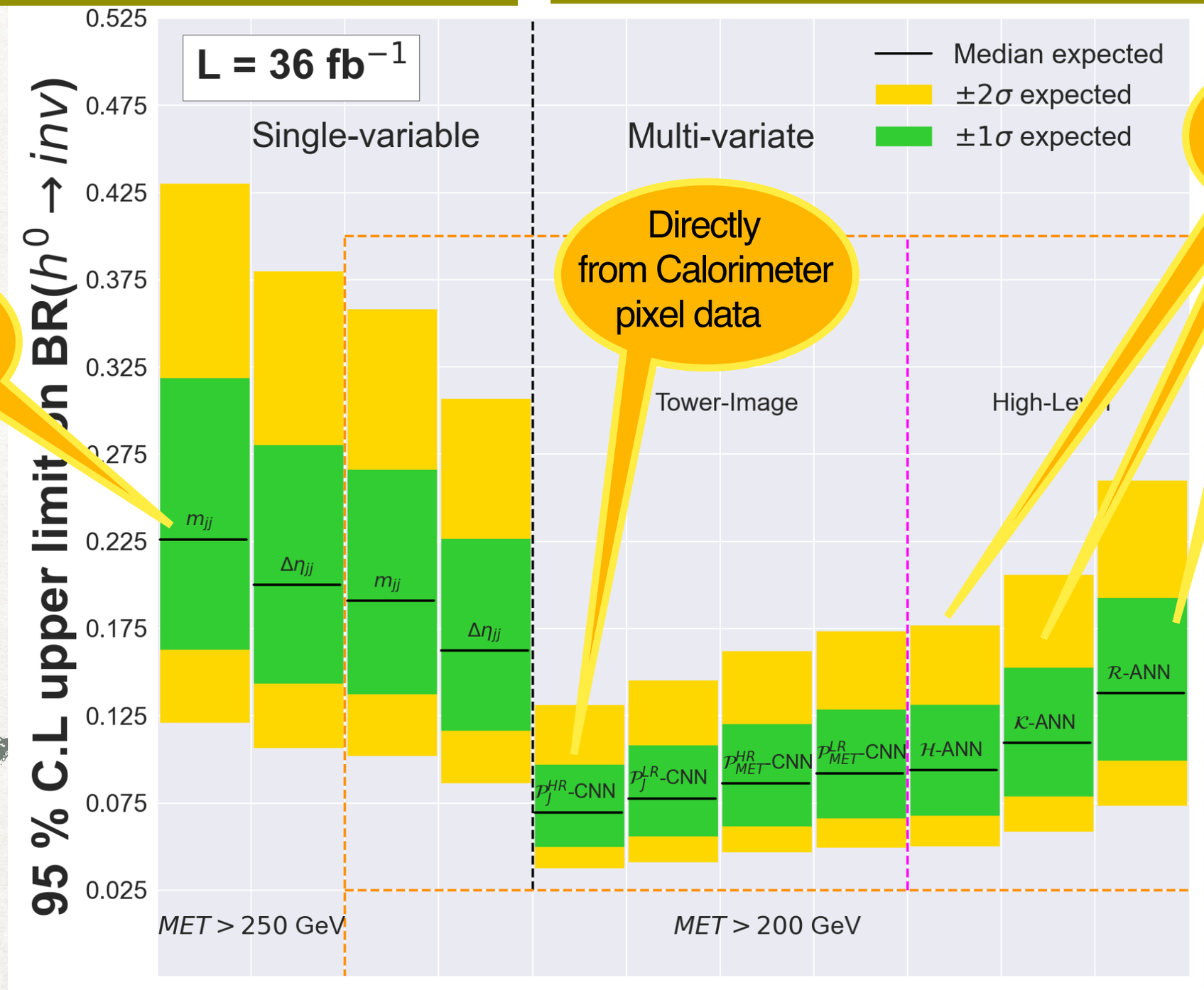
# Deep-learning Invisible Higgs in VBF

Based on: Eur.Phys.J.C 80 (2020) 11, 1055
Phys.Rev.D 105 (2022) 11, 113003

(Akanksha Bhardwaj, Partha Konar, Aruna K Nayak, Vishal Ng)

# INVISIBLE HIGGS DECAY @ VECTOR-BOSON FUSION

*-Based on HL variables constructed by experts-* ○ *—-Based on LL & HL input data———-*




Reproduced CMS analysis result

Directly from Calorimeter pixel data

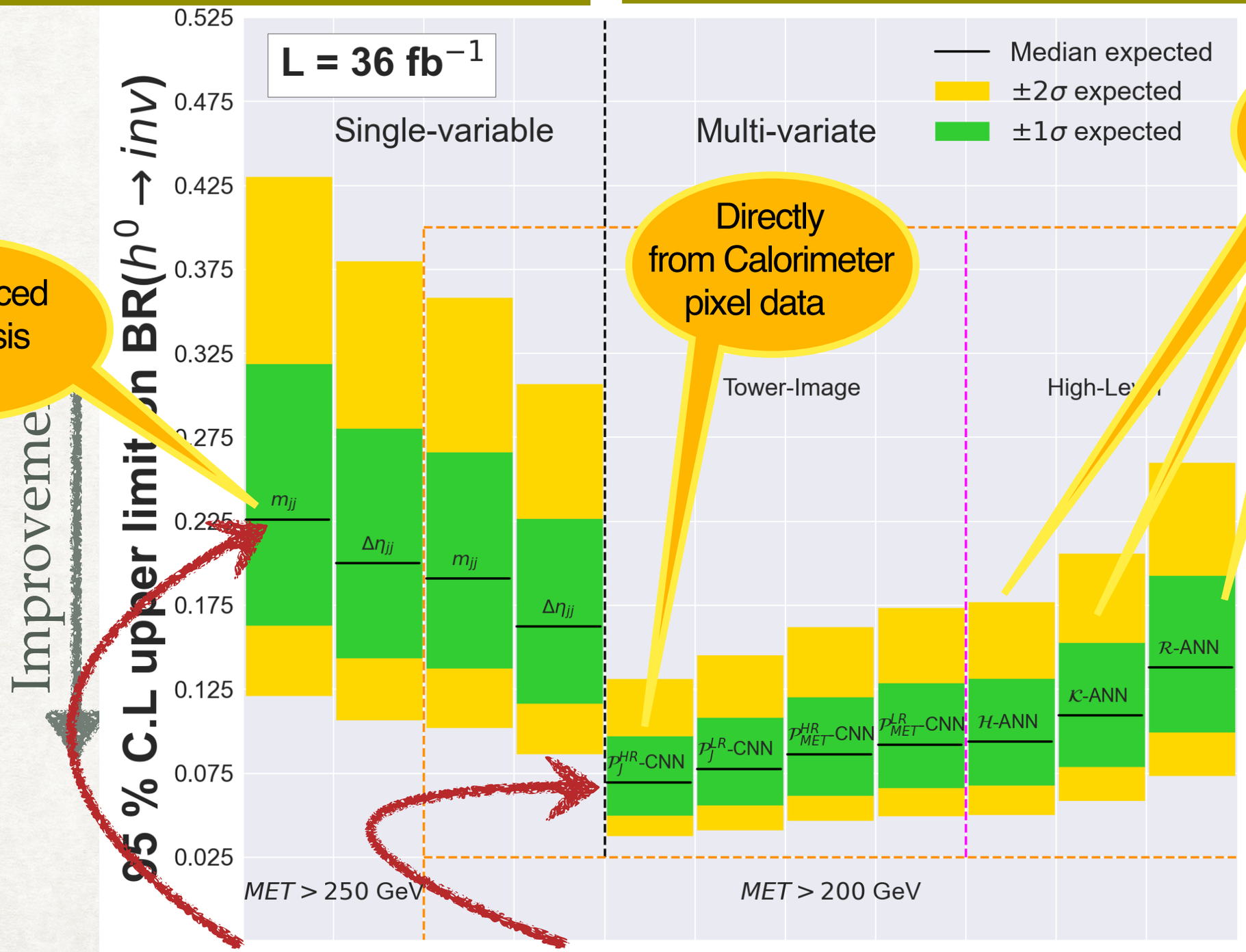
Three High level data analysis

Improvement

*Akanksha Bhardwaj, PK, Aruna Nayak, PK, Vishal Ng; 2022*

# INVISIBLE HIGGS DECAY @ VECTOR-BOSON FUSION

*-Based on HL variables constructed by experts-* ○ *—Based on LL & HL input data———*




Reproduced CMS analysis result

Improvement

Directly from Calorimeter pixel data

Three High level data analysis

Akanksha Bhardwaj, PK, Aruna Nayak, PK, Vishal Ng; 2022

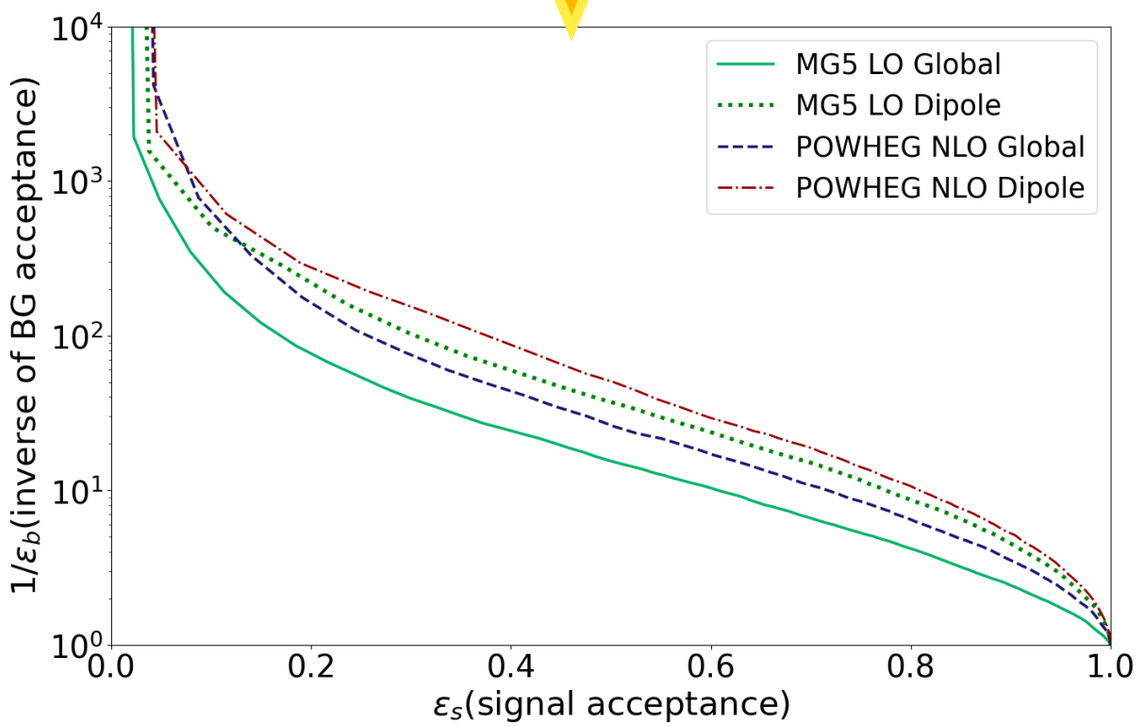
**Factor of three improvement using the same data!**

**Hours of CNN training just extracted the relevant underlying feature better than our decades of research!**

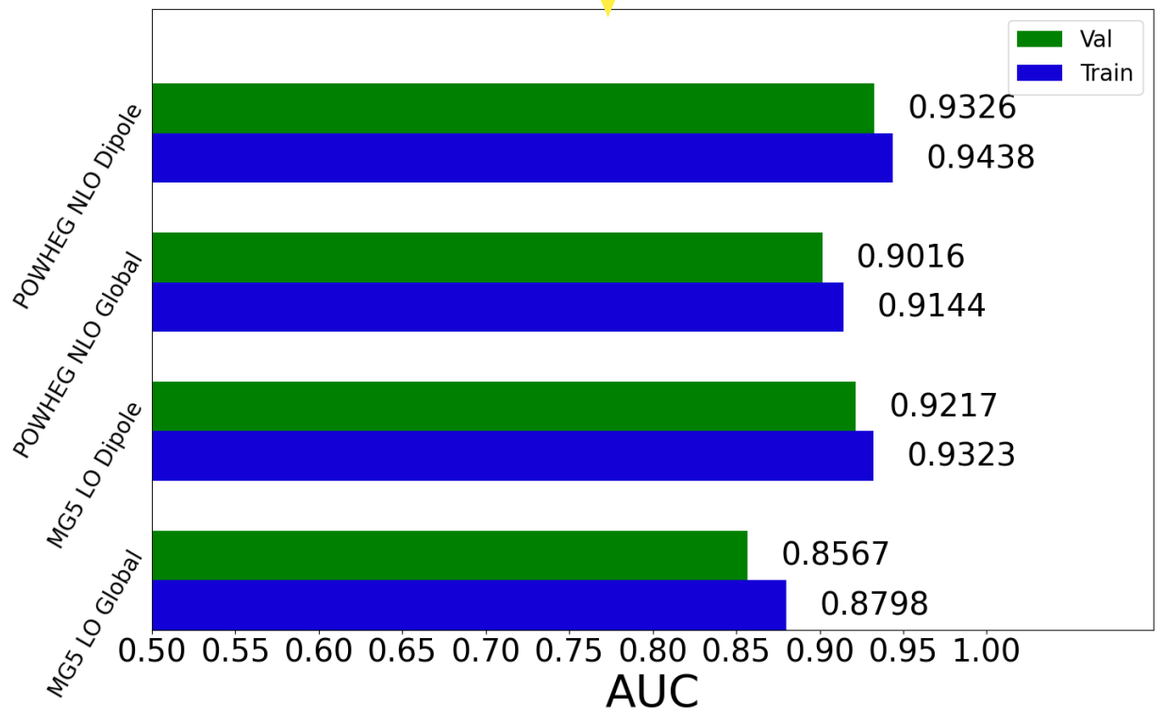
# INVISIBLE HIGGS DECAY @ VECTOR-BOSON FUSION

## ROLE OF PARTON SHOWER

Receiver Operator Characteristics (ROC) Curve




AUC



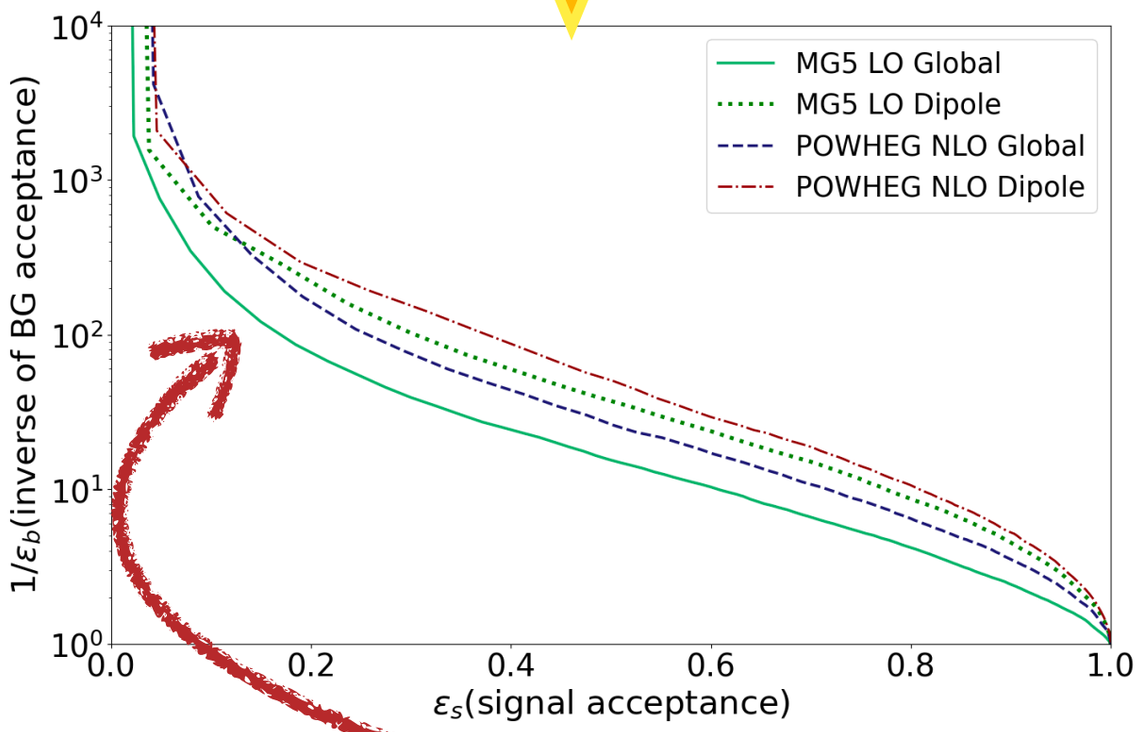

- ✓ LO + Global parton shower scheme shows lowest performance
- ✓ NLO + Dipole parton shower scheme shows best performance
- ✓ Rest two (LO+ Dipole & NLO+ Global) shows intermediate performance

*# Vishal Ng, PK (2021)*

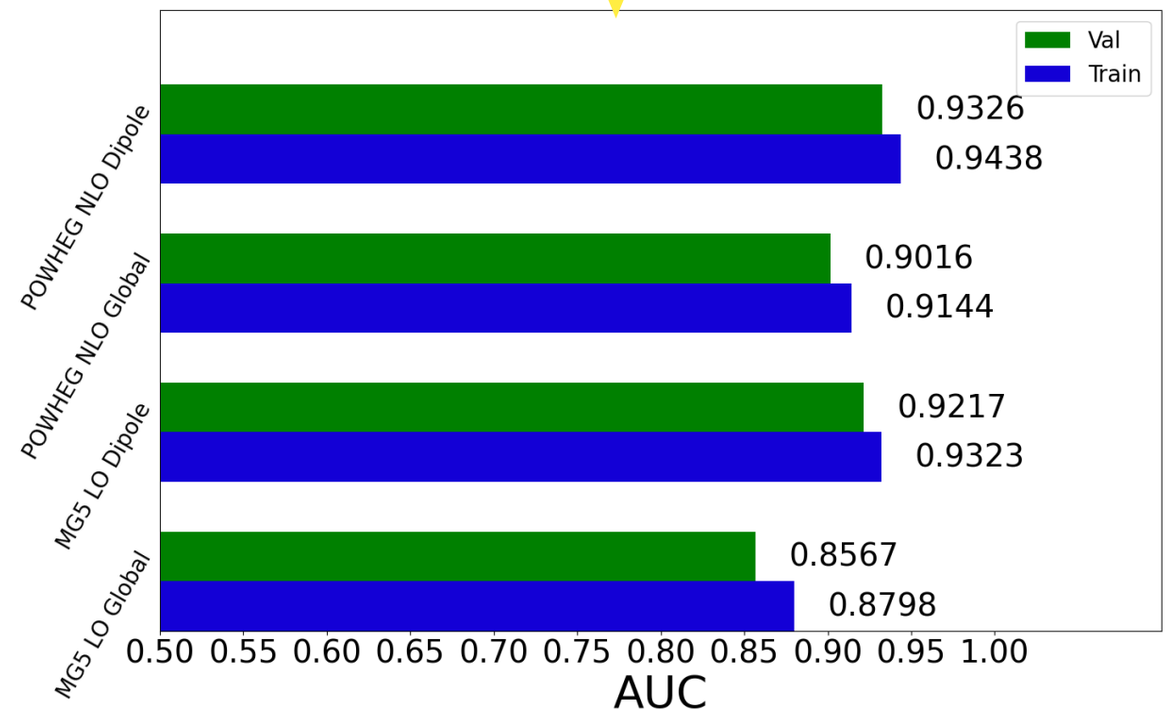
# INVISIBLE HIGGS DECAY @ VECTOR-BOSON FUSION

## ROLE OF PARTON SHOWER

Receiver Operator Characteristics (ROC) Curve




AUC




- ✓ LO + Global parton shower scheme shows lowest performance
- ✓ NLO + Dipole parton shower scheme shows best performance
- ✓ Rest two (LO+ Dipole & NLO+ Global) shows intermediate performance

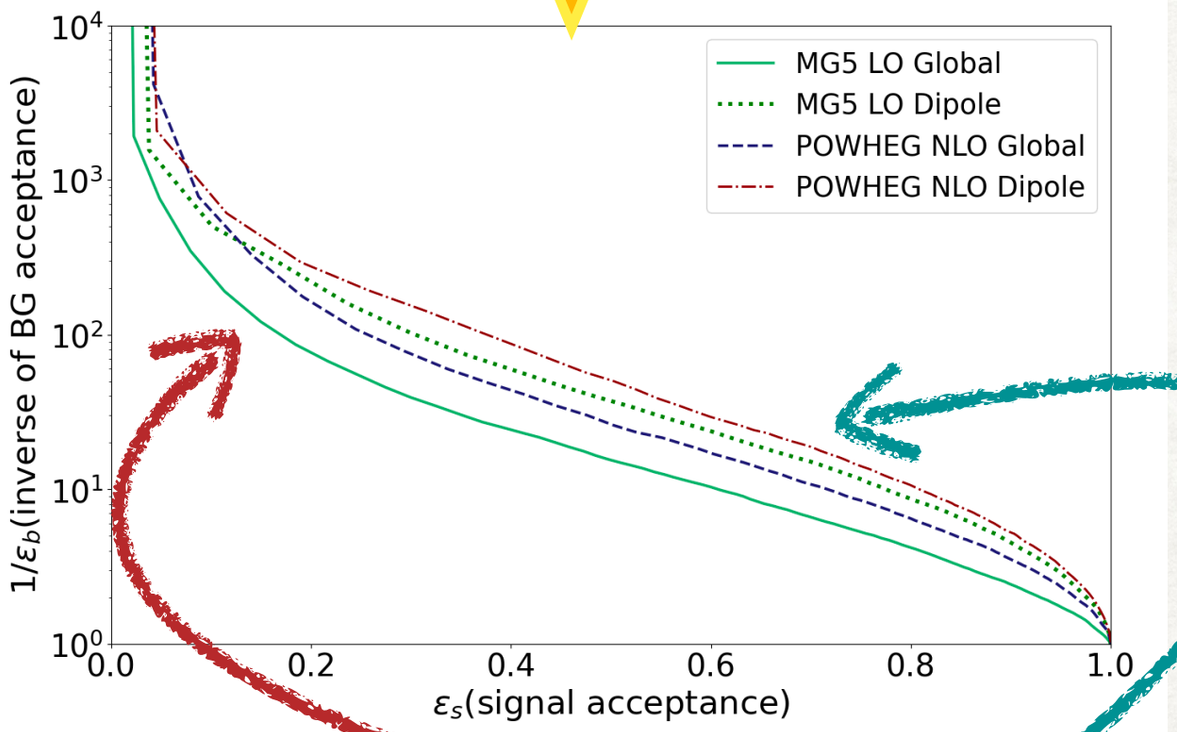
Common in analysis

*# Vishal Ng, PK (2021)*

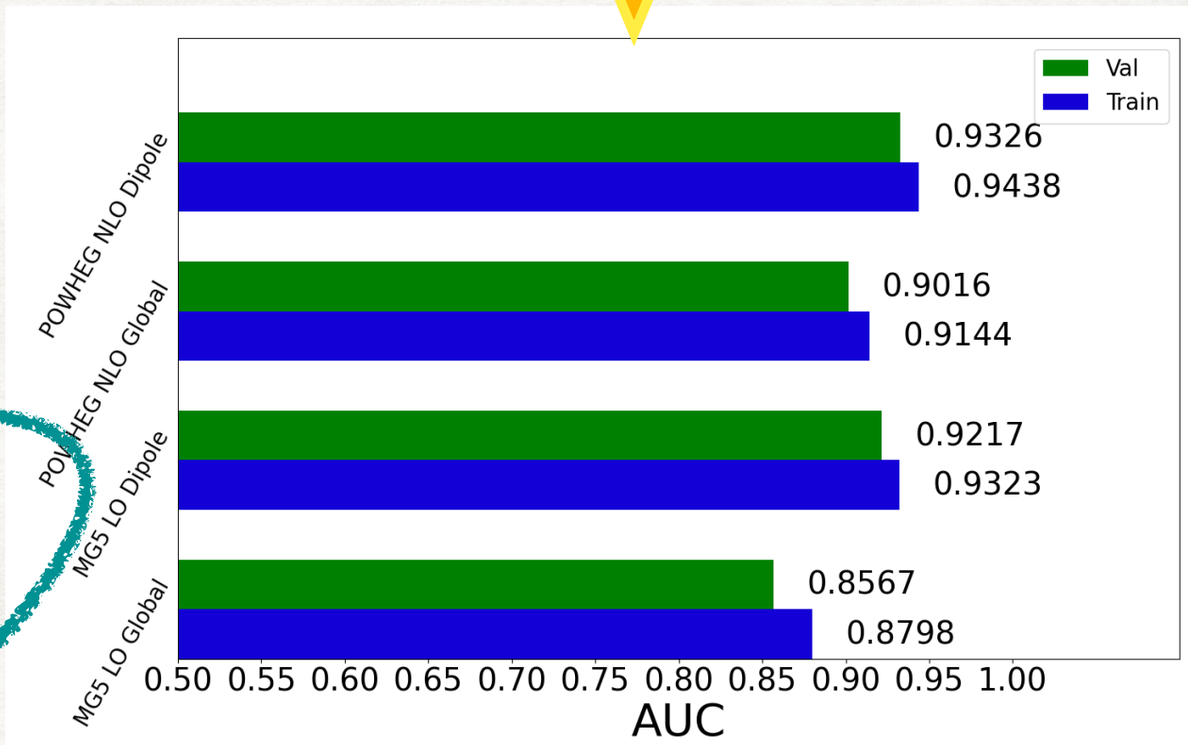
# INVISIBLE HIGGS DECAY @ VECTOR-BOSON FUSION

## ROLE OF PARTON SHOWER

Receiver Operator Characteristics (ROC) Curve




AUC

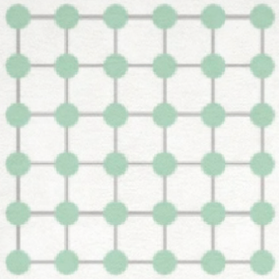



- ✓ LO + Global parton shower scheme shows lowest performance
- ✓ NLO + Dipole parton shower scheme shows best performance
- ✓ Rest two (LO+ Dipole & NLO+ Global) shows intermediate performance

Accurate description

Common in analysis

# Vishal Ng, PK (2021)

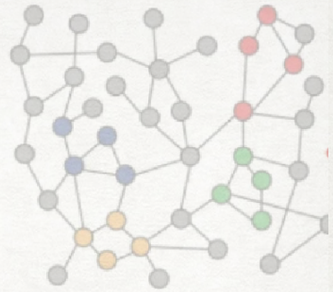


Images

Text

# BEYOND CNN

## GRAPH NEURAL NETWORK

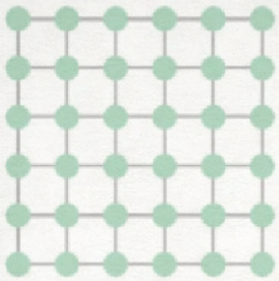


Networks

- Detectors calorimeter hits are typically very sparse and unstructured
- Varying number of reconstructed constituents
- Large number of tunable parameters

✓ Euclidean image (CNN) => general non-Euclidean domain (GNN) : Geometric deep learning

- Graph: Event as point cloud with each entry containing a vector composed of observables
- Graph == Nodes (data point) + Edges (connections are as important as the data itself)
- Message passing operation: nodes features and edge features are exchanged and provide a sophisticated feature extraction
- GNN is very powerful recent concept - mostly unexplored!!



Images

Text

# BEYOND CNN

## GRAPH NEURAL NETWORK

etworks

- Detectors calorimeter hits are typically very spars tructured
- Varying number of reconstructed constitue
- Large number of tunable parameters

✓ Euclidean image (CNN) => g Euclidean domain (GNN) : Geometric deep learning

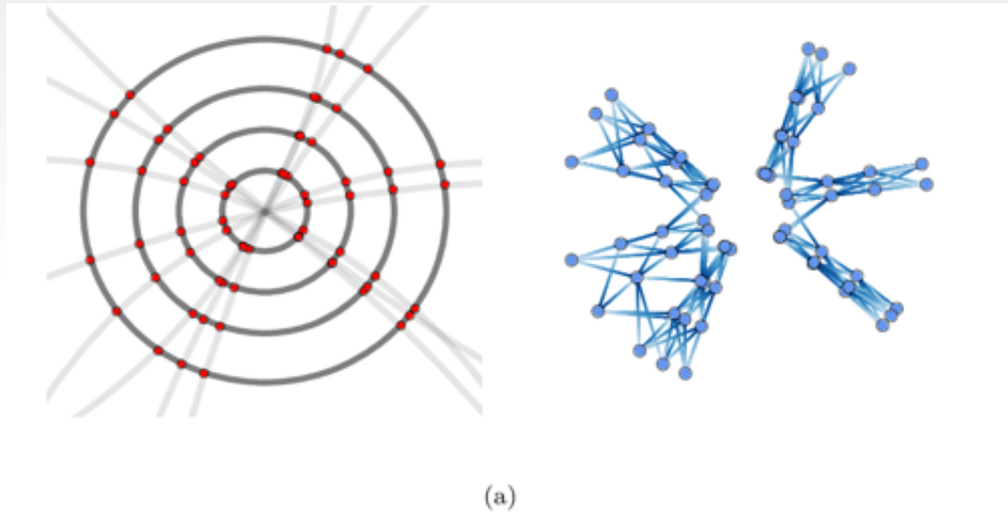
**Heterogeneous data with variable length & Inherently non-Euclidean**
**=> Point clouds provide a flexible geometric representation**

- Graph: Event as with each entry containing a vector composed of
- Graph == data point) + Edges (connections are as important as the
- passing operation: nodes features and edge features are exchanged and provide a sophisticated feature extraction
- GNN is very powerful recent concept - mostly unexplored!!

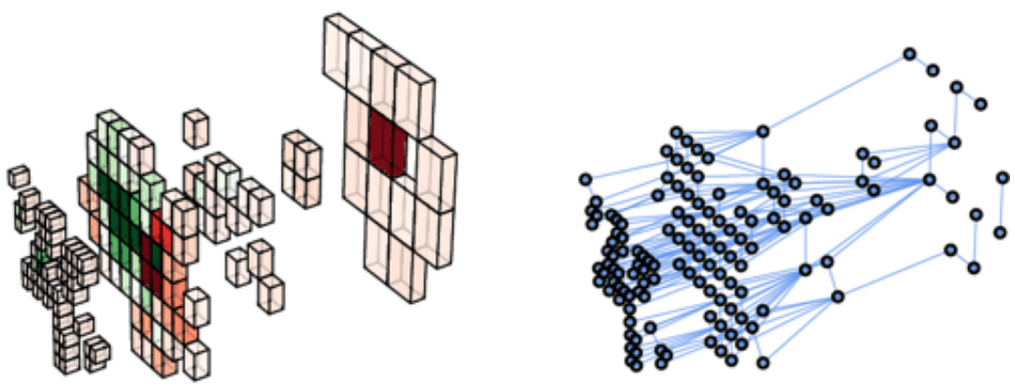
# GNN in HEP

➡ Graph Classification —
Jet Classification : quark/gluon discrimination;
Boosted top/ W/ Z/ Higgs tagging vs QCD jet
Importance of the attention mechanism

➡ Event Classification —
full event identification in collider physics
classification of the signal in the IceCube detector
event classification for other signal topologies



(a)



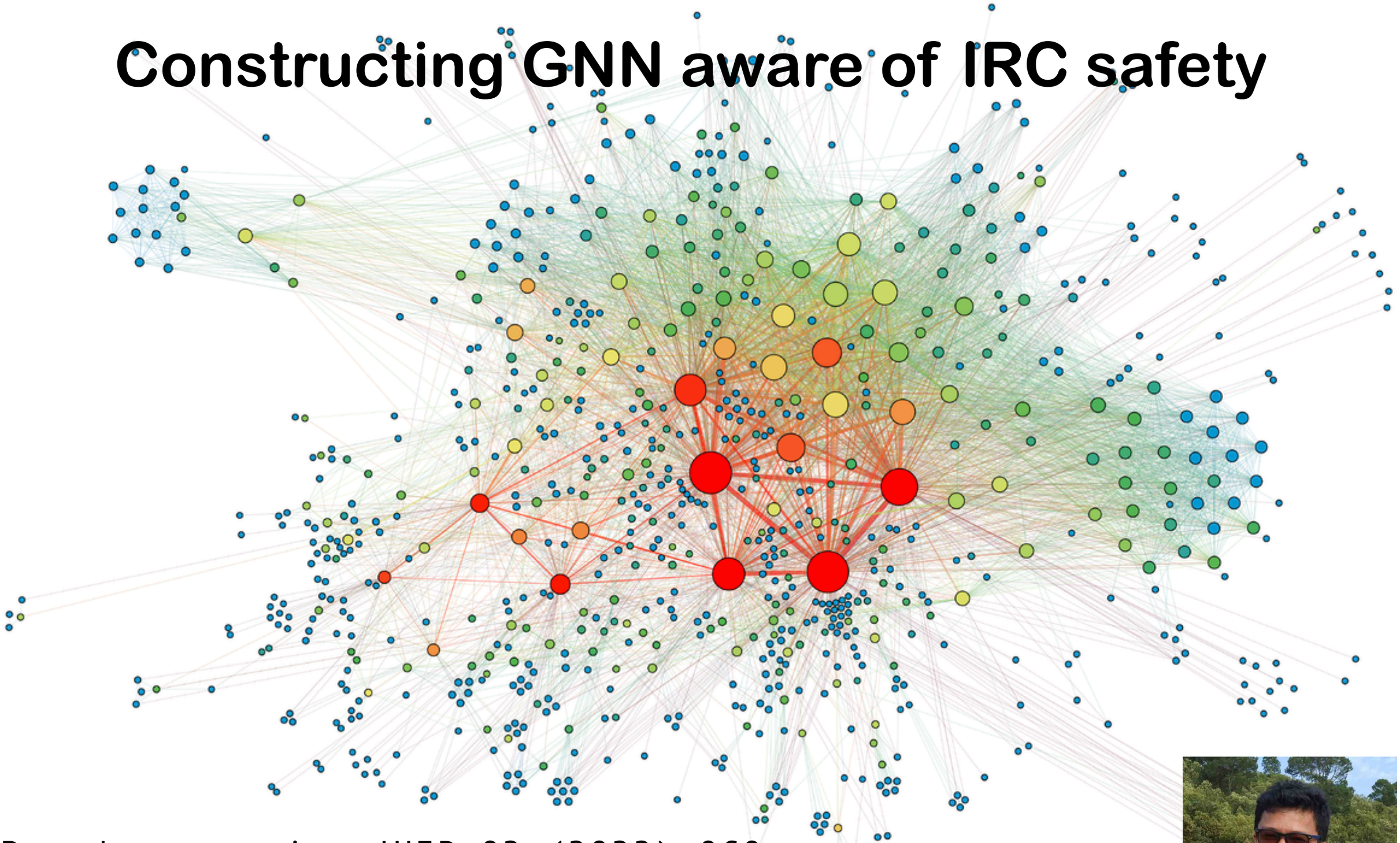
(b)

➡ Node Classification and Regression
Pileup Mitigation
Calorimeter Reconstruction
Particle Flow Reconstruction
per-jet tagging efficiency

➡ Edge Classification
Charged Particle Tracking
Secondary Vertex Reconstruction

2007.13681

# Constructing GNN aware of IRC safety



Based on: arxiv: JHEP 02 (2022) 060

# Cited in newly introduced AI chapter - PDG'22 & 24

JHEP 01 (2024) 113

(Partha Konar, Vishal Ng, Michael Spannowsky)

# CONSTRUCTING PHYSICS AWARE NETWORK

**Any QCD observable should be**

**— sensitive to the physics you want to probe**

**— calculable from first principles in Quantum Chromodynamics (QCD)**

**IRC safety** ensures that the phase space restrictions that the measured value of an observable do not disrupt the cancellation between the real and virtual contributions to the observable at each perturbative order when the soft and collinear regions of phase space are inclusively summed over [Sterman and Weinberg]

# CONSTRUCTING PHYSICS AWARE NETWORK

**Any QCD observable should be**

**— sensitive to the physics you want to probe**

**— calculable from first principles in Quantum Chromodynamics (QCD)**

**IRC safety** ensures that the phase space restrictions that the measured value of an observable do not disrupt the cancellation between the real and virtual contributions to the observable at each perturbative order when the soft and collinear regions of phase space are inclusively summed over [Sterman and Weinberg]

**How can we make neural networks aware of this physics input? So that, it treats all hadronic/jet analysis in a IRC safe way.**

# CONSTRUCTING PHYSICS AWARE NETWORK

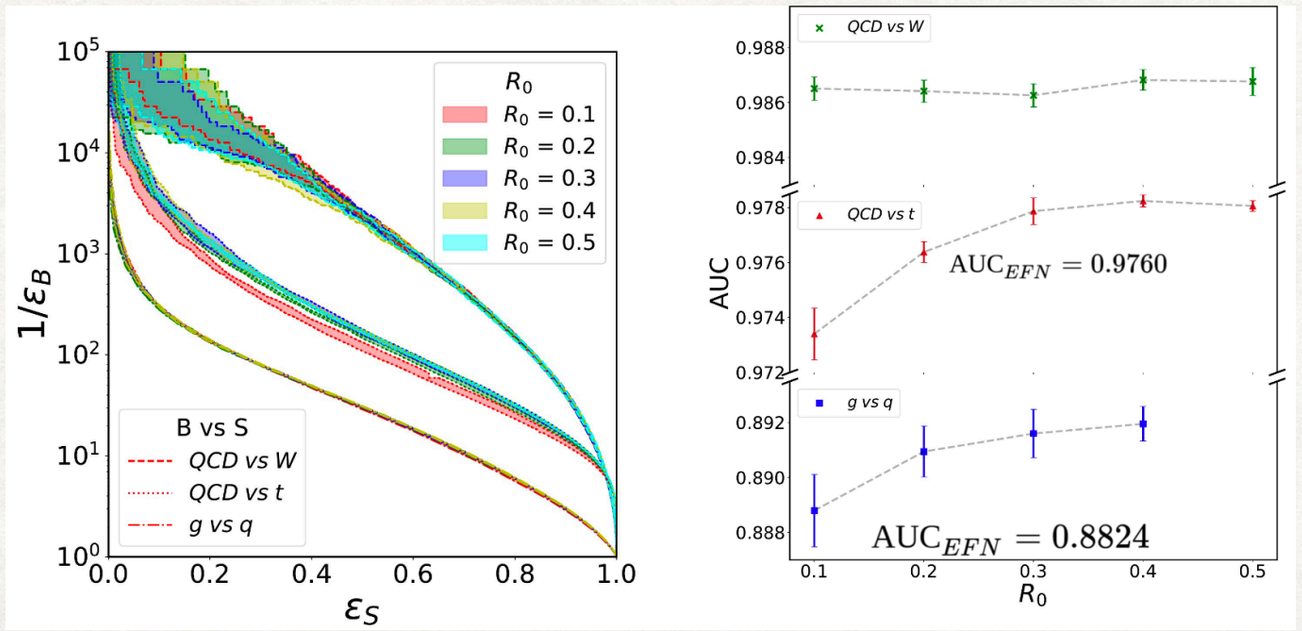
**Any QCD observable should be**

**— sensitive to the physics you want to probe**

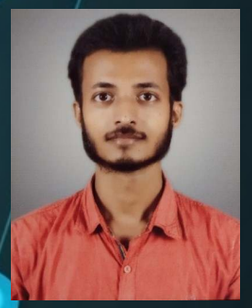
**— calculable from first principles in Quantum Chromodynamics (QCD)**

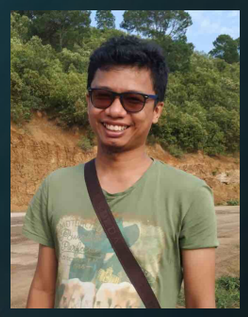
**IRC safety** ensures that the phase space restrictions that the measured value of an observable do not disrupt the cancellation between the real and virtual contributions to the observable at each perturbative order when the soft and collinear regions of phase space are inclusively summed over [Sterman and Weinberg]

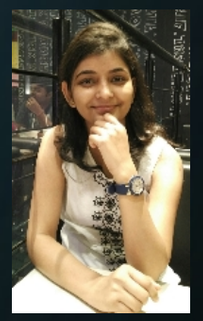
**How can we make neural networks aware of this physics input? So that, it treats all hadronic/jet analysis in a IRC safe way.**



# Interpretability of Deep Learning: — Opening the Black Box

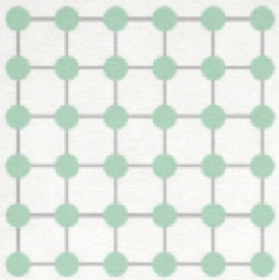






e-Print: 2404.16207 and follow up with Deepanshu

(Akanksha Bhardwaj, Partha Konar, Vishal Ng)





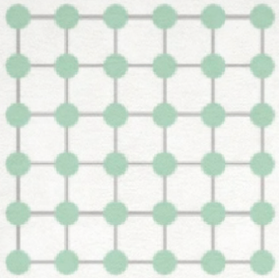
# WHY INTERPRETABILITY MATTERS

## THE INNER WORKINGS OF BLACK BOX MODELS


- Difficult to understand how they arrive at their decisions
- Lack of transparency can hinder trust, debugging
- Limit deployment of AI systems ensuring fairness, accountability, and safety

- As researcher, often we look for understanding the laws of nature, rather than efficiency!

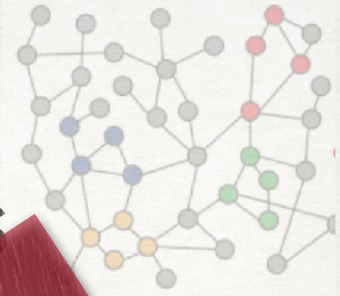
- Dependence, Dominance, effective analytic expressions
- Symmetry and Group
- ..





# WHY INTERPRETABILITY MATTERS

## THE INNER WORKINGS OF BLACK BOX MODELS

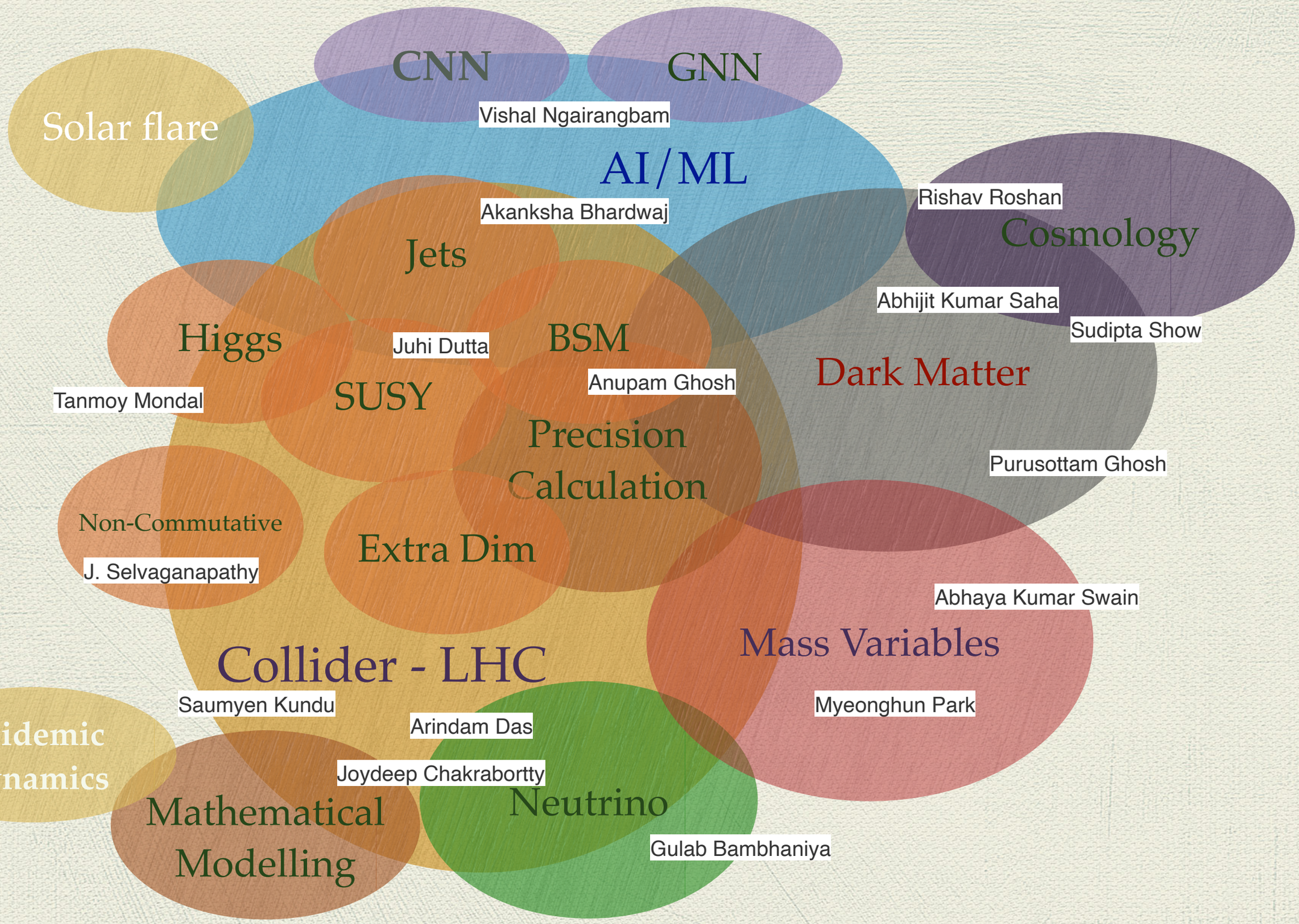


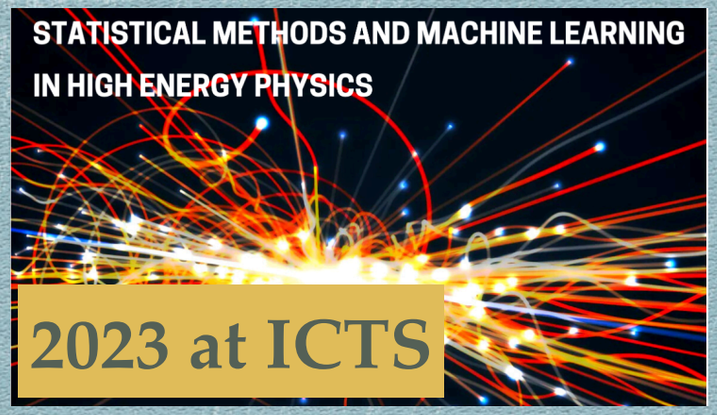

- Difficult to understand how they arrive at their decisions
- Lack of transparency can hinder trust, debugging
- Limit deployment of AI systems ensuring fairness, accountability, and safety

XML: Yet unexplored highly desirable

- As researcher, often we look for understanding the laws of nature, rather than efficiency!

- Dependence, Dominance, effective analytic expressions
- Symmetry and Group
- ..


CNN
GNN
Vishal Ngairangbam
Solar flare
AI/ML
Akanksha Bhardwaj
Rishav Roshan
Cosmology
Jets
Abhijit Kumar Saha
Sudipta Show
Higgs
Juhi Dutta
BSM
Dark Matter
Anupam Ghosh
Tanmoy Mondal
SUSY
Precision
Calculation
Purusottam Ghosh
Non-Commutative
J. Selvaganapathy
Extra Dim
Abhaya Kumar Swain
Mass Variables
Collider - LHC
Myeonghun Park
Saumyen Kundu
Arindam Das
Epidemic
Dynamics
Joydeep Chakrabortty
Mathematical
Modelling
Neutrino
Gulab Bambhaniya




Primarily for Students and PDFs working on using deep learning

Preparatory school (Online)
[June 12 - 23 2023]

Lecture + Tutorial
28 Aug - 04 Sep 2023

Workshop
5 - 8 Sep 2023

https://www.icts.res.in/program/ml4hep

ML4HEP
IOPB 2024

2024 at IOP

# Machine Learning for Particle and Astroparticle Physics

ML4HEP 2024