



# DATA SCIENCE



VSKP DSS 2020-21 BATCH PRESENTING YOU



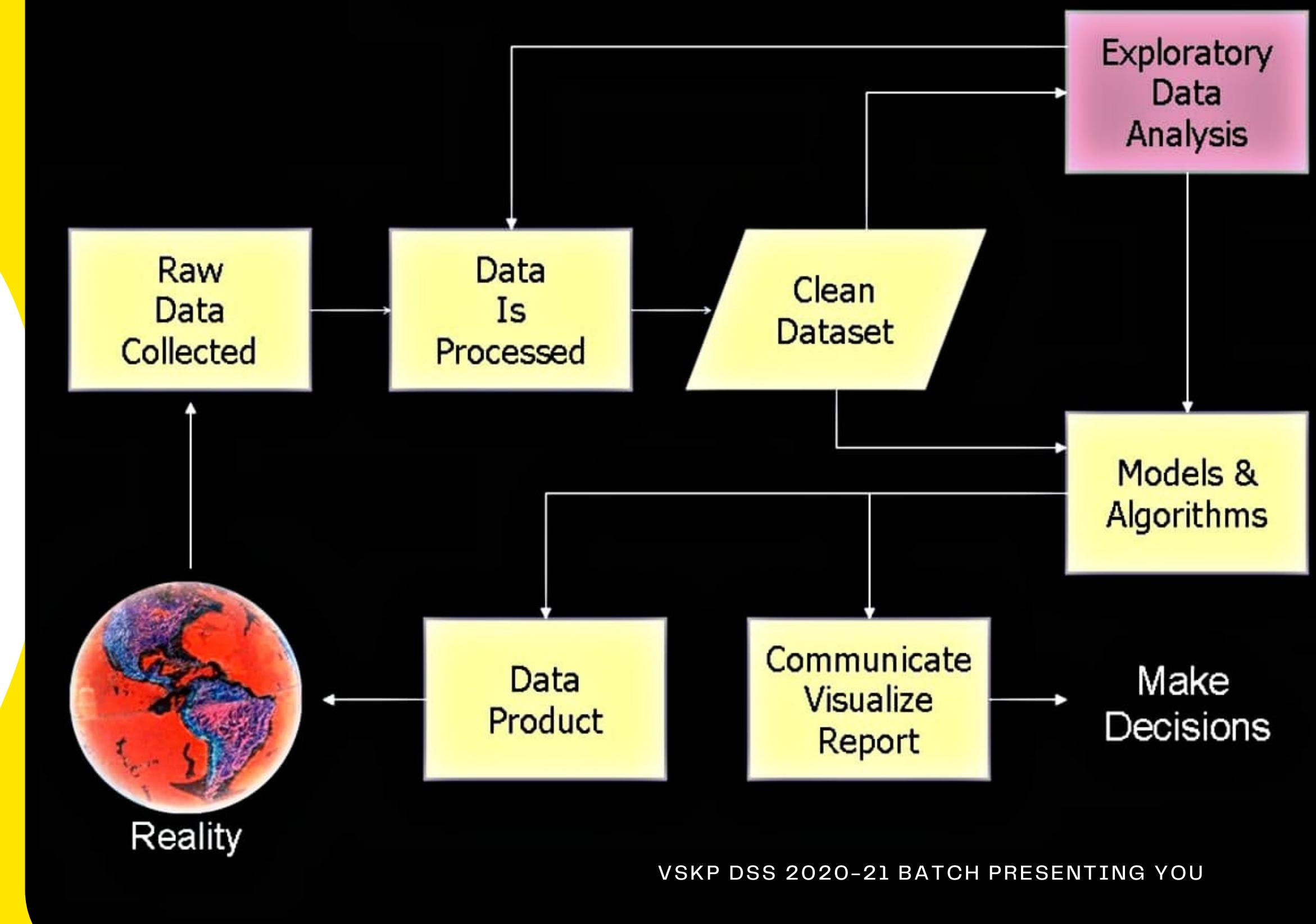
# **What is Data Science?**



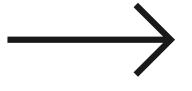
**DATA SCIENCE IS THE FIELD OF STUDY THAT COMBINES DOMAIN EXPERTISE,  
PROGRAMMING SKILLS, AND KNOWLEDGE OF MATHEMATICS AND STATISTICS TO  
EXTRACT MEANINGFUL INSIGHTS FROM DATA.**

# Data Science Process and Terminologies

- 1 COLLECTION OF RAW DATA**
- 2 DATA CLEANING**
- 3 EXPLORATORY DATA ANALYSIS**
- 4 MODEL & ALGORITHMS**
- 5 DATA VISUALIZATION**



# DATA SCIENCE SCOPE & CAREERS



## **What does Data Analyst do?**

Data analysts swifts through data and seek to identify trends

## **What does Data Scientist do?**

Data scientists are pros at interpreting data, but also tend to have coding and mathematical modeling expertise.

## **What type of Companies need people in Data Science?**

Every company needs people in Data Science, As every company has raw data that should be converted into meaningful data & using the trends the company can make objective decisions for scaling up the business.

## **What makes it unique?**

It serves a variety of purposes, making data interpretation powerful tools for visualization and prediction.



# Why is it the Hot Cake ?



- 1. DATA SCIENCE HAS QUIETLY GROWN TO INCLUDE BUSINESSES AND ORGANIZATIONS WORLD-WIDE**
- 2. IT HAS BECOME AN IMPORTANT PART OF BUSINESS AND ACADEMIC RESEARCH**
- 3. DATA SCIENCE REVOLUTION HAS BEEN A GRADUAL SHIFT TO WRITING MORE AND MORE CONSERVATIVE PROGRAMMING**
- 4. “INNOVATIVE” CHANGES HAPPEN LESS AND LESS OFTEN**

# Applications of Data Science



Targeted Advertising

Health care industry

Drug development

Airline route planning

Auto complete

and many more..





# Prominent Researchers



**Dhanurjay Patil**  
US Cheif Data Scientist

Dhanurjay Patil is a former US Chief Data Scientist, and along with Jeff Hammerbacher he coined the term “data



**Alex “Sandy” Pentland**  
Co-Lead of WE Forum

co-leads the WE Forum Big & Personal Data initiatives, Founder of the Advisory Boards for Motorola Mobility, and a variety of start-ups



**Dean Abbott**  
Data Sicienctist

Founder and president of Abbott Analytics. With over 21 years of enriching experience in data preparation and data visualization.



TOP RESEARCHER THAT STANDS OUT OF ALL

# Andrew Yan-Tak Ng



A PIONEER IN ONLINE EDUCATION





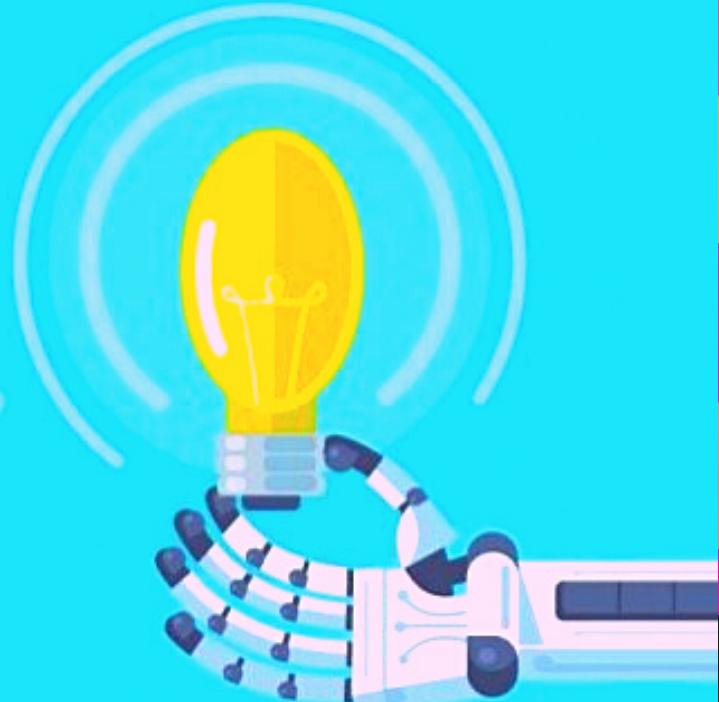
# Who is he?

## A Brief About Him:

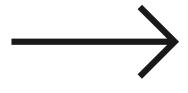
- British-born American businessman
- Computer Scientist
- Professor at Stanford University
- Co-founded Coursera and deeplearning.ai
- How brands work online

“AI IS THE NEW  
ELECTRICITY.”

ANDREW NG



# What he Believes?



## About AI

AI is the new  
electricity

## Industry

It is difficult to think of a major industry that AI will not transform. This includes healthcare, education, transportation, retail, communications, and agriculture. There are surprisingly clear paths for AI to make a big difference in all of these industries.

## Future Vision

Elon Musk is worried about AI apocalypse, but I am worried about people losing their jobs. The society will have to adapt to a situation where people learn throughout their lives depending on the skills needed in the marketplace.

## What makes him unique

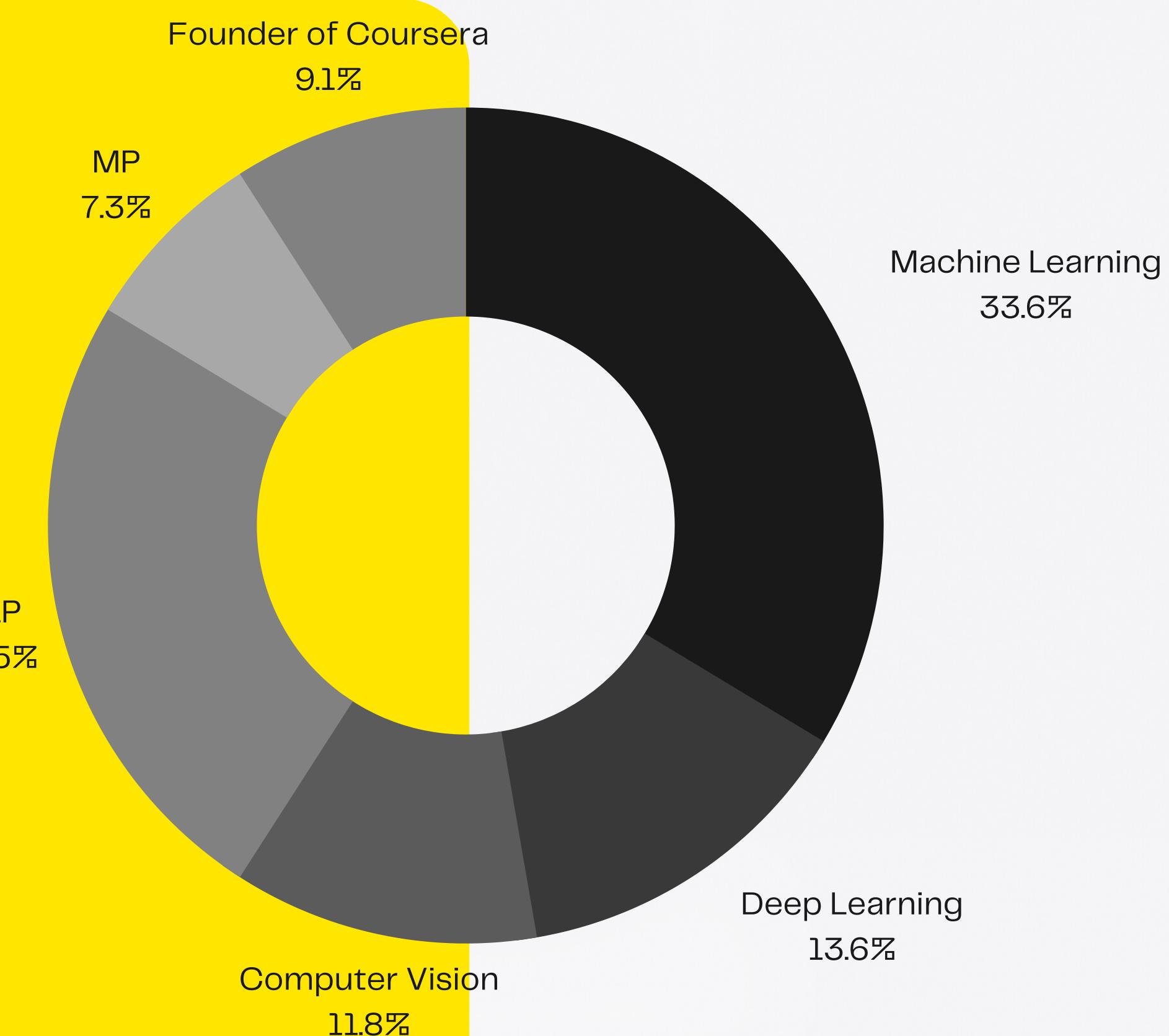
A novel approach to walk you through a real-life project and the associated problems.



# What does he do?

## FIELDS OF RESEARCH

Ng researches primarily in machine learning, deep learning, machine perception, computer vision, and natural language processing; and is one of the world's most famous and influential computer scientists.



# Andrew's Top Attributes & Work



**developed one of the most capable autonomous helicopters in the world**

**Worked on the STAIR (STanford Artificial Intelligence Robot) project, which resulted in ROS, robotics software platform.**

**Founding team members for the Stanford WordNet project, which uses ML to expand the Princeton WordNet database created by Christiane Fellbaum**

**Founded the Google Brain project at Google, which developed large-scale artificial neural networks using Google's distributed computer infrastructure**

**Co-authored the influential paper that introduced Latent Dirichlet allocation**

**Founding of Coursera with Koller in 2012**

VSKP DSS 2020-2021 BATCH PRESENTING YOU

# Geoffrey Everest Hinton

- British-Canadian computer scientist
- AI Pioneer
- Received the 2018 Turing Award
- Together with Yoshua Bengio and Yann LeCun, they are called "Godfathers of AI" and "Godfathers of Deep Learning"



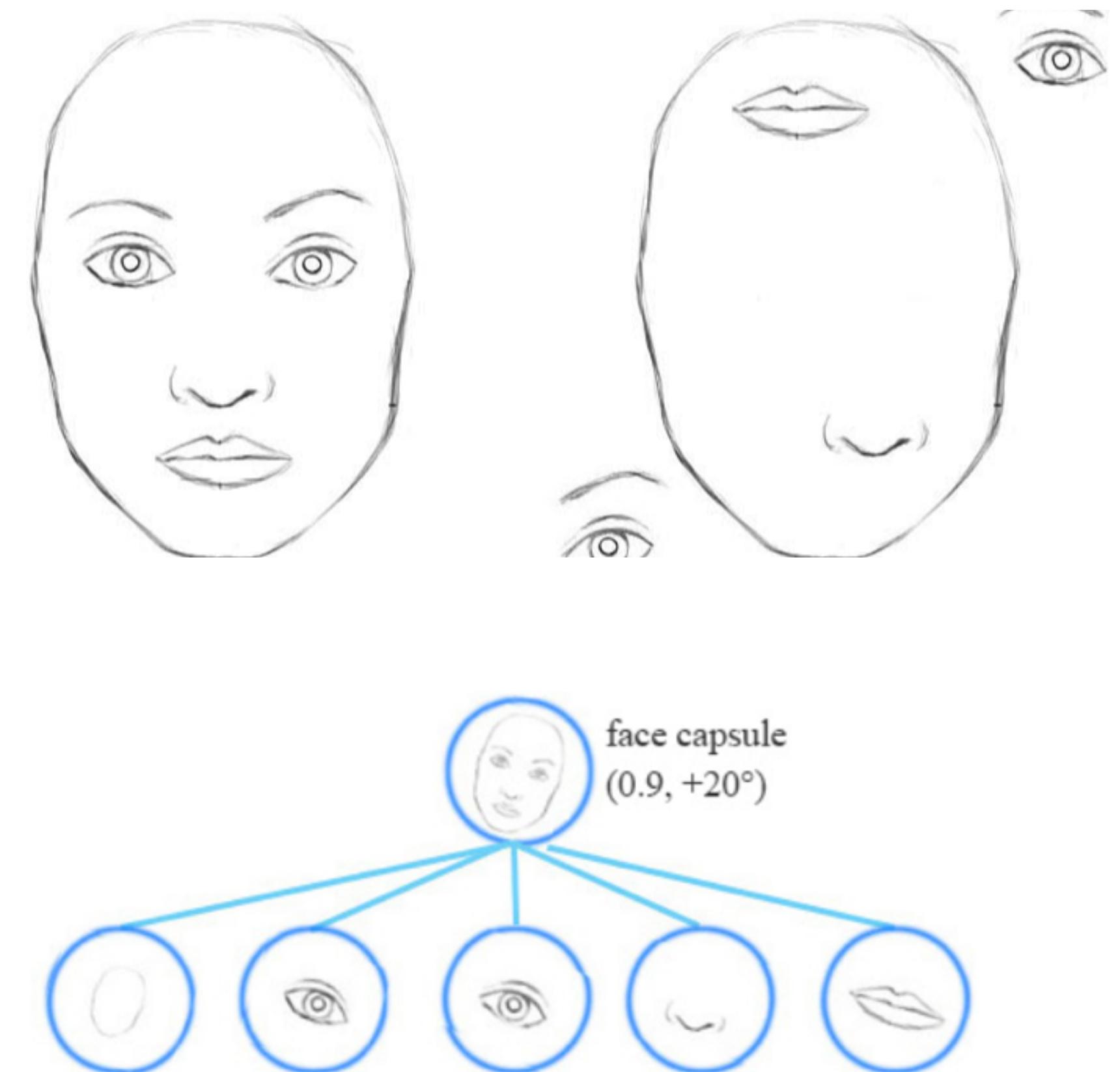
**Geoffrey Hinton:  
“Deep learning is  
going to be able  
to do everything”**

## **Research Work:**

- Designs machine learning algorithms.
- Backpropagation Algorithm.
- Neural networks
- Boltzmann Machine
- Capsule Neural Network
- Deep learning

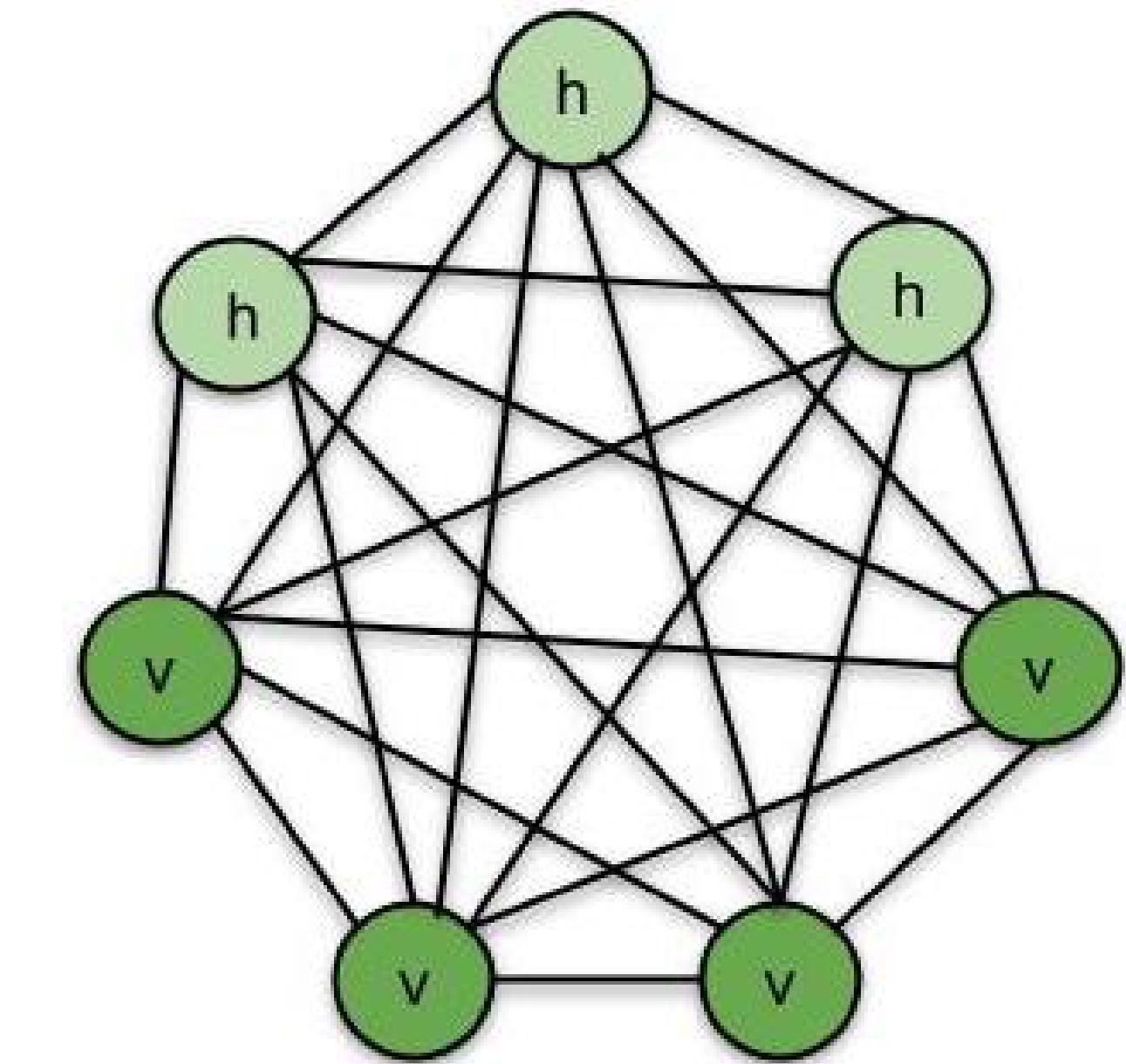
# Capsule Neural Network (CapsNet)

- Models hierarchical relationships.
- The idea is to add structures called "capsules" to a convolutional neural network (CNN)
- Addresses the "Picasso problem" in image recognition.
- In Hinton's original idea one minicolumn would represent and detect one multidimensional entity.



# BOLTZMANN MACHINE

- Boltzmann Machine was invented by Geoffrey Hinton and Terry Sejnowski in 1985.
- It is an unsupervised DL model in which every node is connected to every other node.
- The main purpose of Boltzmann Machine is to optimize the solution of a problem.
- It is the work of Boltzmann Machine to optimize the weights and quantity related to that particular problem.



v - visible nodes, h - hidden nodes