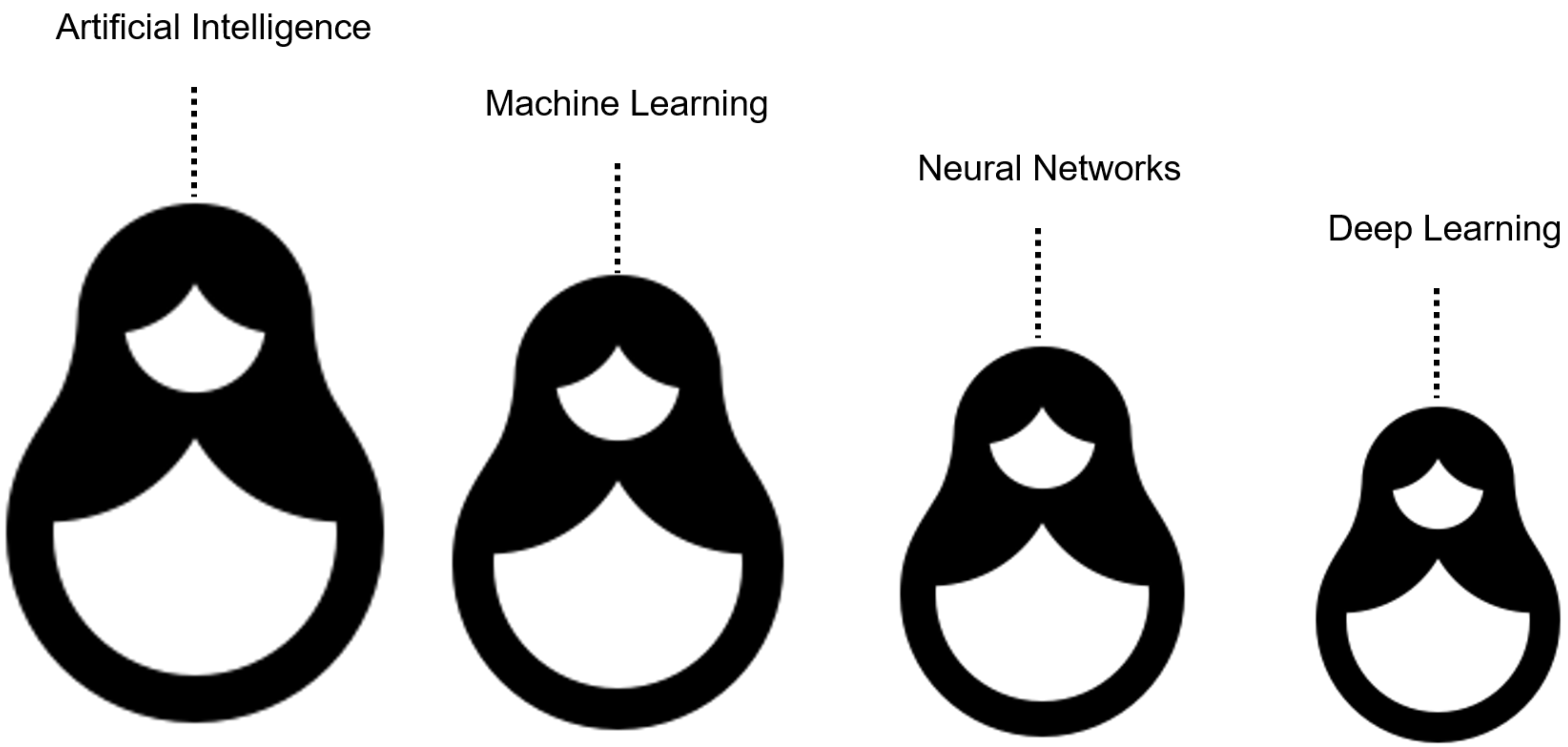
Machine Learning & Deep Learning

Chun-Hsiang Chan
Undergraduate Program in Intelligent Computing and Big Data, Chung Yuan Christian University
Master Program in Intelligent Computing and Big Data, Chung Yuan Christian University

Outline

- Artificial intelligence
- Machine learning
- Deep learning
- Machine learning & deep learning
- Optimization
- Potential issues
- Paper reading
- Assignment

Artificial intelligence



https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks

Machine learning – Origin

https://www.telecomtv.com/content/network-automation/the-human-side-of-network-automation-45860/

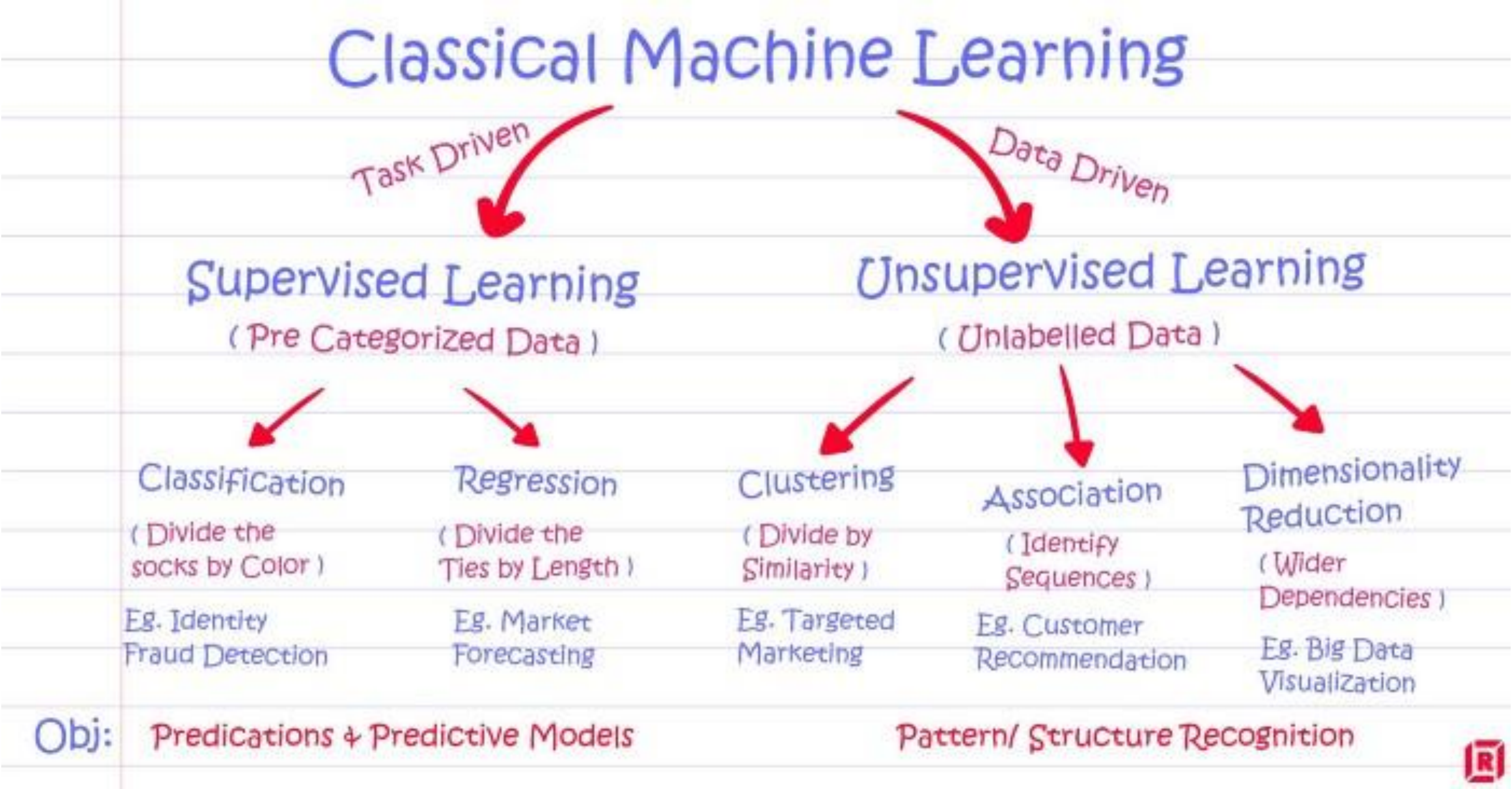
Machine learning – Definition

- **Machine learning** (**ML**) is a field of inquiry devoted to understanding and building methods that "learn" – that is, methods that leverage data to improve performance on some set of tasks. It is seen as a part of artificial intelligence.
- Machine learning algorithms build a model based on sample data, known as training data, in order to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of applications, such as in medicine, email filtering, speech recognition, agriculture, and computer vision, where it is difficult or unfeasible to develop conventional algorithms to perform the needed tasks.

Machine learning – Category

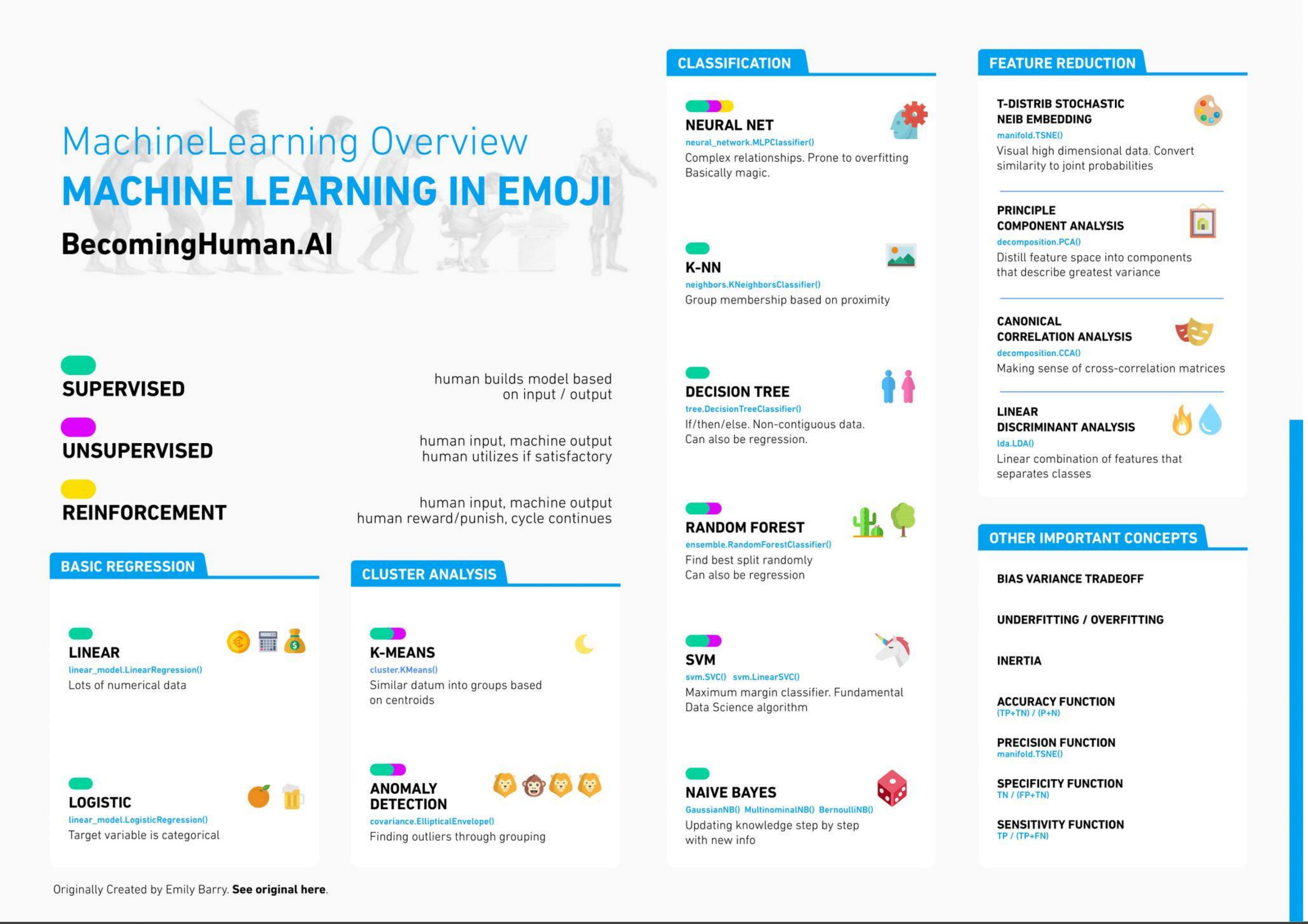
- A subset of machine learning is closely related to computational statistics, which focuses on making predictions using computers, but not all machine learning is statistical learning. The study of mathematical optimization delivers methods, theory and application domains to the field of machine learning. Data mining is a related field of study, focusing on exploratory data analysis through unsupervised learning.

Supervised and unsupervised learning



https://medium.com/@dkatzman_3920/supervised-vs-unsupervised-learning-and-use-cases-for-each-8b9cc3ebd301

Machine learning



Machine learning

Cheat-Sheet Skicit learn
Phyton For Data Science

BecomingHuman.AI DataCamp

Skicit Learn

Skicit Learn is an open source Phyton library that implements a range if machine learning, processing, cross validation and visualization algorithm using a unified

A basic Example

```
>>> from sklearn import neighbors, datasets, preprocessing
>>> from sklearn.cross validation import train_test_split
>>> from sklearn.metrics import accuracy_score
>>> iris = datasets.load _iris() >>> X, y = iris.data[:, :2], iris.target
>>> Xtrain, X test, y_train, y test = train_test_split (X, y, random stat33)
>>> scaler = preprocessing.StandardScaler().fit(X_train)
>>> X train = scaler.transform(X train)
>>> X test = scaler.transform(X test)
>>> knn = neighbors.KNeighborsClassifier(n_neighbors=5)
>>> knn.fit(X_train, y_train)
>>> y_pred = knn.predict(X_test)
>>> accuracy_score(y_test, y_pred)
```

Prediction

Supervised Estimators

```
>>> y_pred = svc.predict(np.random.radom((2,5)))     Predict labels
>>> y_pred = lr.predict(X_test)                       Predict labels
>>> y_pred = knn.predict_proba(X_test)                Estimate probability of a label
```

Unsupervised Estimators

```
>>> y_pred = k_means.predict(X_test)                  Predict labels in clustering algos
```

Loading the Data

Your data beeds to be nmueric and stored as NumPy arrays or SciPy sparse matric. other types that they are comvertible to numeric arrays, such as Pandas Dataframe, are also acceptable

```
>>> import numpy as np >> X = np.random.random((10,5))
>>> y = np . array ( PH', IM', 'F', 'F' , 'M', 'F', 'NI', 'tvl' , 'F', 'F', 'F' ))
>>> X [X < 0.7] = 0
```

Preprocessing The Data

Standardization

```
>>> from sklearn.preprocessing import StandardScaler
>>> scaler = StandardScaler().fit(X_train)
>>> standardized_X = scaler.transform(X_train)
>>> standardized_X_test = scaler.transform(X_test)
```

Normalization

```
>>> from sklearn.preprocessing import Normalizer
>>> scaler = Normalizer().fit(X_train)
>>> normalized_X = scaler.transform(X_train)
>>> normalized_X_test = scaler.transform(X_test)
```

Binarization

```
>>> from sklearn.preprocessing import Binarizer
>>> binarizer = Binarizer(threshold=0.0).fit(X)
>>> binary_X = binarizer.transform(X)
```

Encoding Categorical Features

```
>>> from sklearn.preprocessing import Imputer
>>> imp = Imputer(missing_values=0, strategy='mean', axis=0)
>>> imp.fit_transform(X_train)
```

Imputing Missing Values

```
>>> from sklearn.preprocessing import Imputer
>>> imp = Imputer(missing_values=0, strategy='mean', axis=0)
>>> imp.fit_transform(X_train)
```

Generating Polynomial Features

```
>>> from sklearn.preprocessing import PolynomialFeatures
>>> poly = PolynomialFeatures(5)
>>> poly.fit_transform(X)
```

Evaluate Your Model's Performance

Classification Metrics

Accuracy Score

```
>>> knn.score(X_test, y_test)                          Estimator score method
>>> from sklearn.metrics import accuracy_score         Metric scoring functions
>>> accuracy_score(y_test, y_pred)
```

Classification Report

```
>>> from sklearn.metrics import classification_report  Precision, recall, f1-score
>>> print(classification_report(y_test, y_pred))       and support
```

Confusion Matrix

```
>>> from sklearn.metrics import confusion_matrix
>>> print(confusion_matrix(y_test, y_pred))
```

Regression Metrics

Mean Absolute Error

```
>>> from sklearn.metrics import mean_absolute_error
>>> y_true = [3, -0.5, 2]
>>> mean_absolute_error(y_true, y_pred)
```

Mean Squared Error

```
>>> from sklearn.metrics import mean_squared_error
>>> mean_squared_error(y_test, y_pred)
```

R^2 Score

```
>>> from sklearn.metrics import r2_score
>>> r2_score(y_true, y_pred)
```

Clustering Metrics

Adjusted Rand Index

```
>>> from sklearn.metrics import adjusted_rand_score
>>> adjusted_rand_score(y_true, y_pred)
```

Homogeneity

```
>>> from sklearn.metrics import homogeneity_score
>>> homogeneity_score(y_true, y_pred)
```

V-measure

```
>>> from sklearn.metrics import v_measure_score
>>> metrics.v_measure_score(y_true, y_pred)
```

Cross-Validation

```
>>> from sklearn.cross_validation import cross_val_score
>>> print(cross_val_score(knn, X_train, y_train, cv=4))
>>> print(cross_val_score(lr, X, y, cv=2))
```

Model Fitting

Supervised learning

```
>>> lr.fit(X, y)                                       Fit the model to the data
>>> knn.fit(X_train, y_train)
>>> svc.fit(X_train, y_train)
```

Unsupervised Learning

```
>>> k_means.fit(X_train)                               Fit the model to the data
>>> pca_model = pca.fit_transform(X_train)             Fit to data, then transform it
```

Create Your Model

Supervised Learning Estimators

Linear Regression

```
>>> from sklearn.linear_model import LinearRegression
>>> lr = LinearRegression(normalize=True)
```

Support Vector Machines (SVM)

```
>>> from sklearn.svm import SVC
>>> svc = SVC(kernel='linear')
```

Naive Bayes

```
>>> from sklearn.naive_bayes import GaussianNB
>>> gnb = GaussianNB()
```

KNN

```
>>> from sklearn import neighbors
>>> knn = neighbors.KNeighborsClassifier(n_neighbors=5)
```

Unsupervised Learning Estimators

Principal Component Analysis (PCA)

```
>>> from sklearn.decomposition import PCA
>>> pca = PCA(n_components=0.95)
```

K Means

```
>>> from sklearn.cluster import KMeans
>>> k_means = KMeans(n_clusters=3, random_state=0)
```

Training And Test Data

```
>> from sklearn.cross validation import train_test_split
>> X train, X test, y train, y test - train_test_split(X,
                                                        y,
                                                        random state-0)
```

Tune Your Model

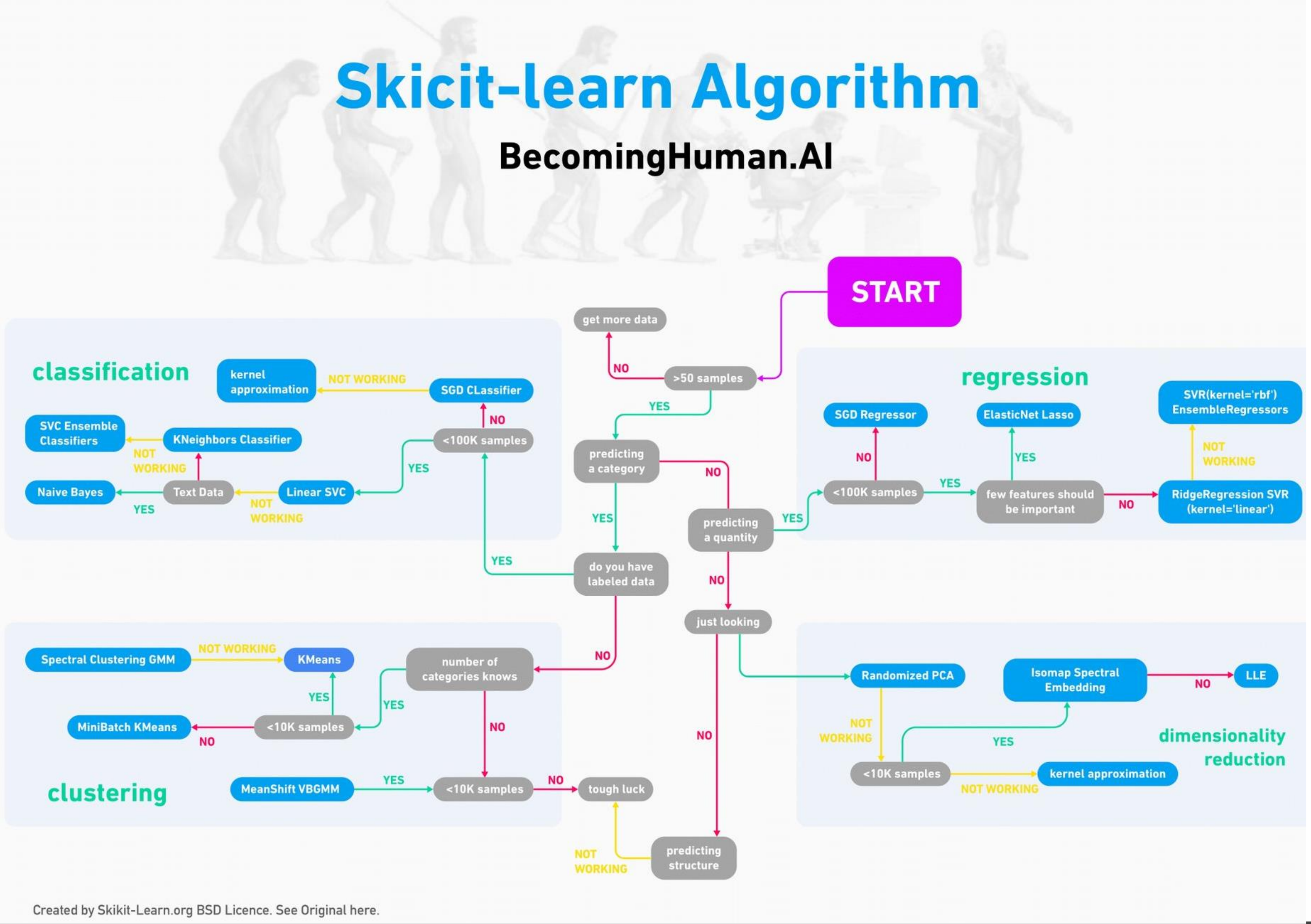
Grid Search

```
>>> from sklearn.grid_search import GridSearchCV
>>> params = {"n_neighbors": np.arange(1,3)
              "metric": ["euclidean","cityblock"]}
>>> grid = GridSearchCV(estimator=knn,
                        param_grid=params)
>>> grid.fit(X_train, y_train)
>>> print(grid.best_score_)
>>> print(grid.best_estimator_.n_neighbors)
```

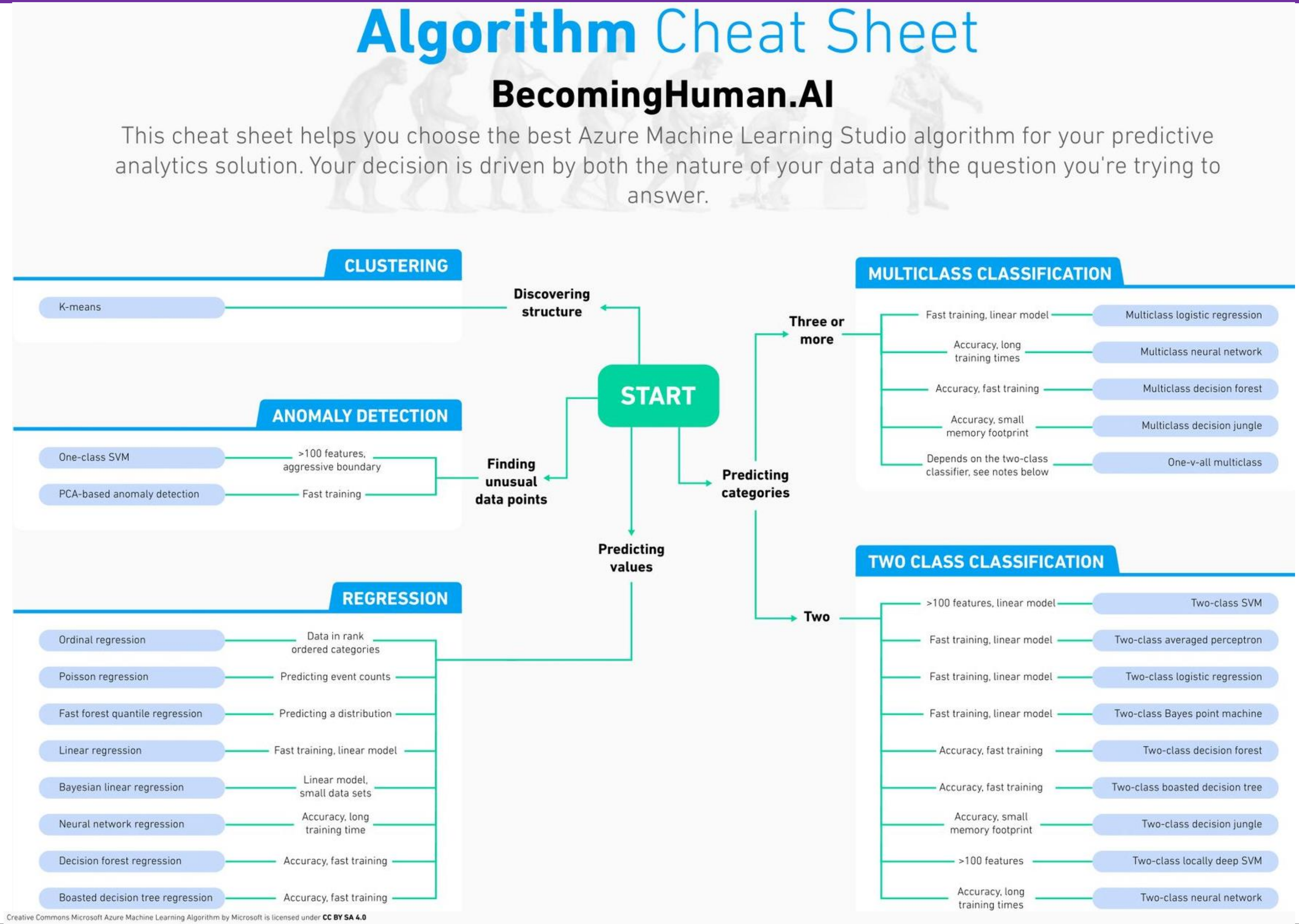
Randomized Parameter Optimization

```
>>> from sklearn.grid_search import RandomizedSearchCV
>>> params = {"n_neighbors": range(1,5),
              "weights": ["uniform", "distance"]}
>>> rsearch = RandomizedSearchCV(estimator=knn,
                                 param_distributions=param
                                 cv=4,
                                 n_iter=8,
                                 random_state=5)
>>> rsearch.fit(X_train, y_train)
>>> print(rsearch.best_score_)
```

Machine learning

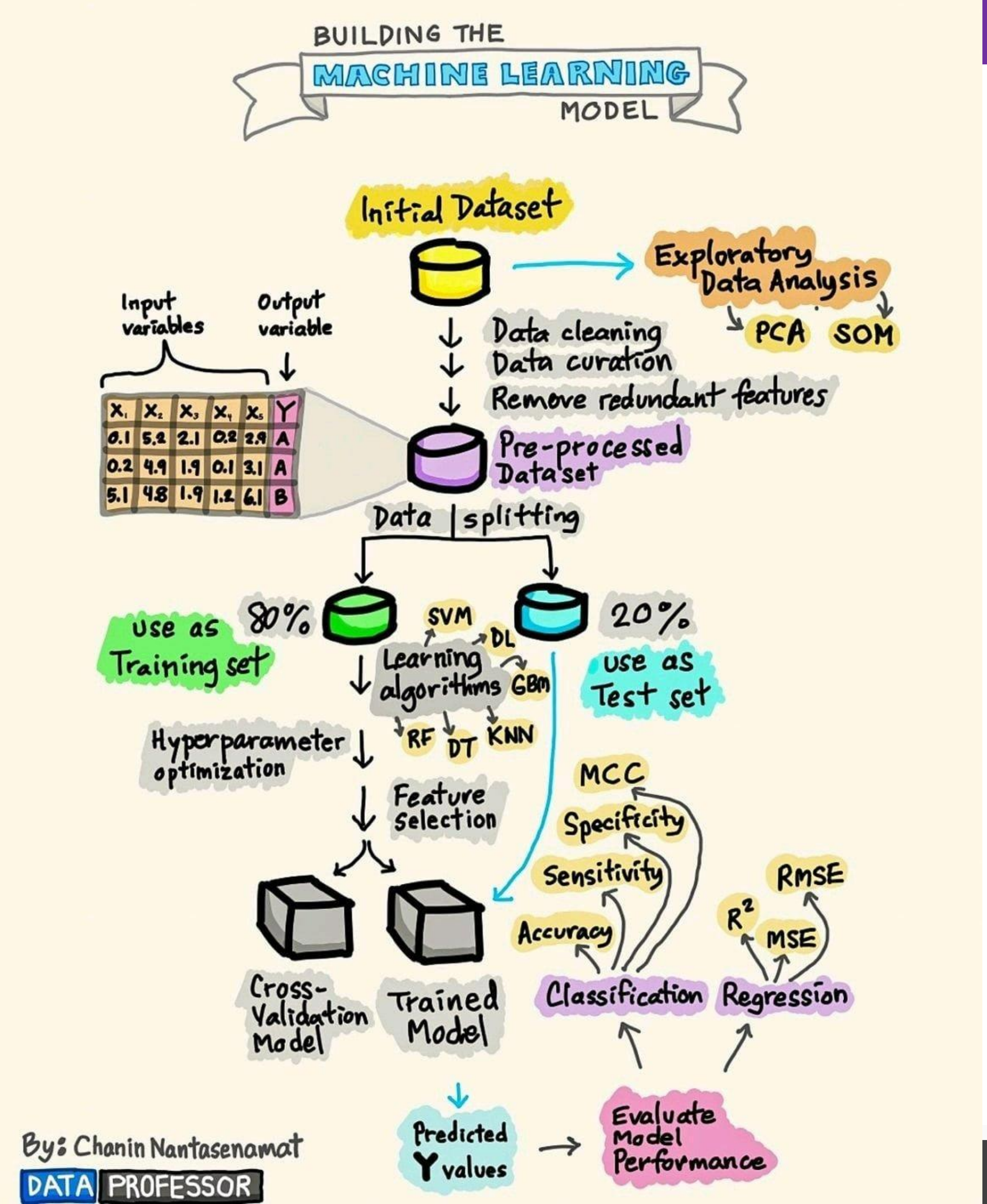


Machine learning



Machine learning
Evaluation metrics

To measure the performance of the model



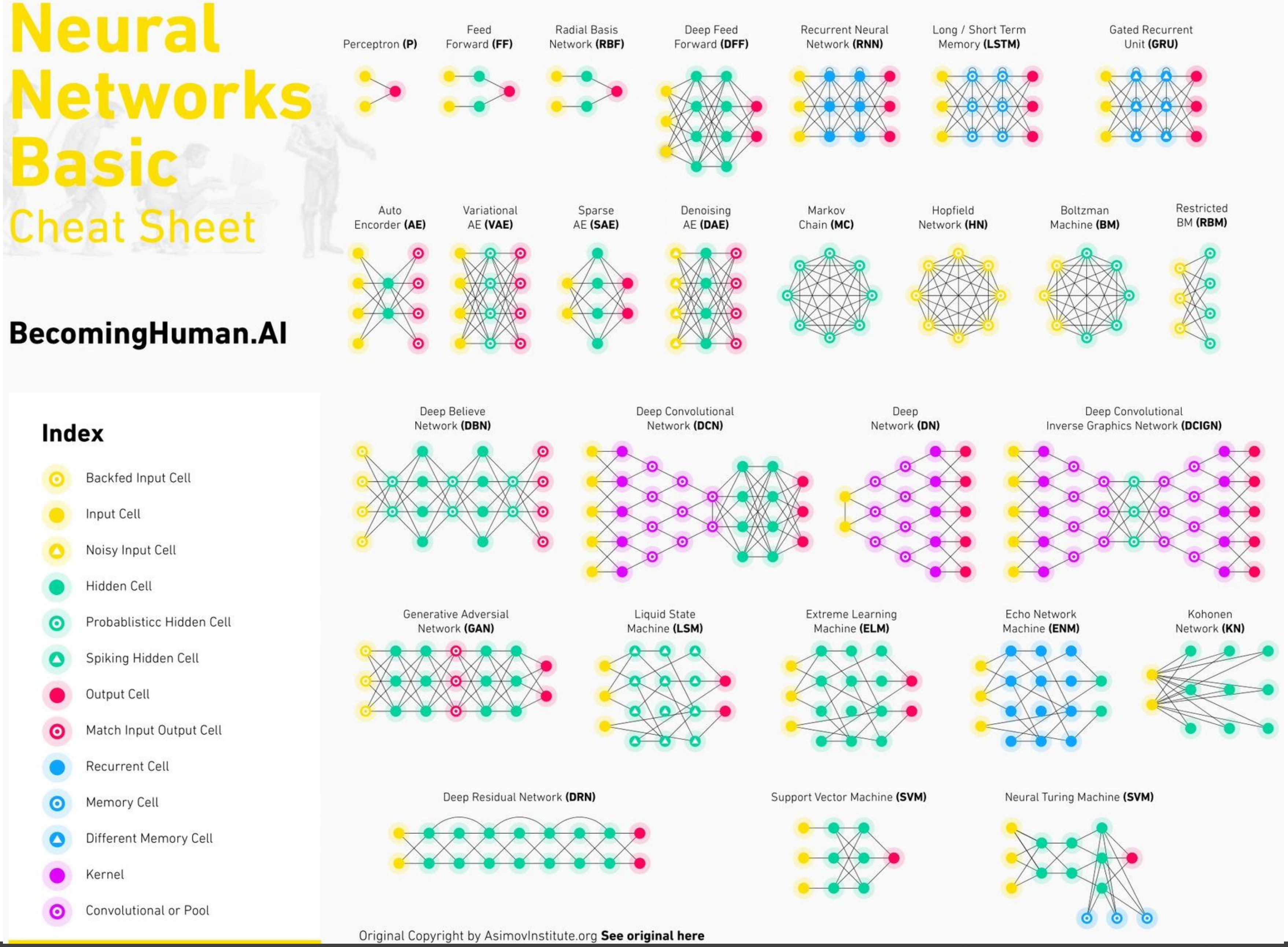
Deep learning

- **Deep learning (DL)** neural networks, or artificial neural networks, attempts to mimic the human brain through a combination of data inputs, weights, and bias. These elements work together to accurately recognize, classify, and describe objects within the data.
- Deep neural networks consist of multiple layers of interconnected nodes, each building upon the previous layer to refine and optimize the prediction or categorization. This progression of computations through the network is called forward propagation. The input and output layers of a deep neural network are called *visible* layers. The input layer is where the deep learning model ingests the data for processing, and the output layer is where the final prediction or classification is made.

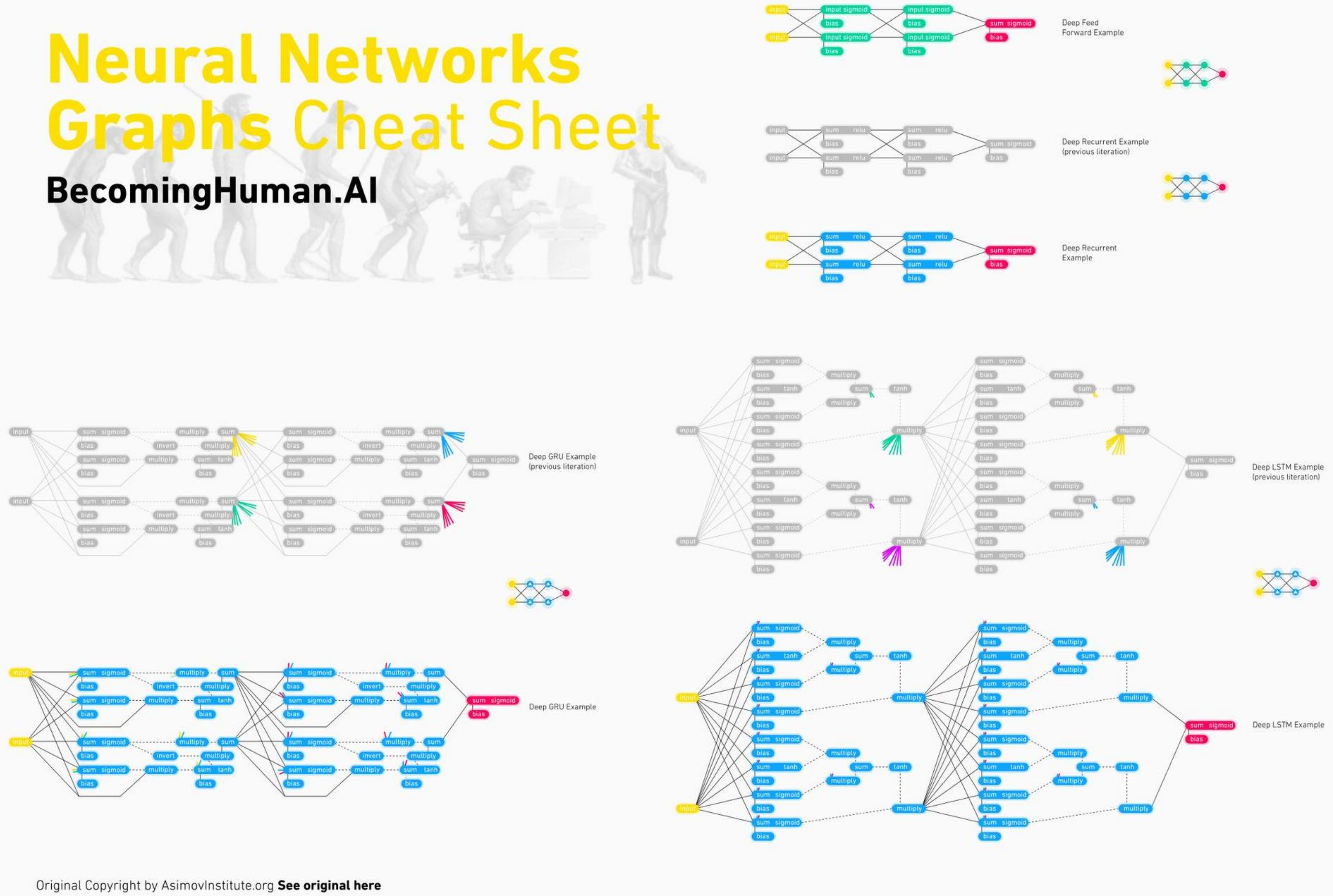
Deep learning

- Another process called backpropagation uses algorithms, like gradient descent, to calculate errors in predictions and then adjusts the weights and biases of the function by moving backwards through the layers in an effort to train the model.
- Together, forward propagation and backpropagation allow a neural network to make predictions and correct for any errors accordingly. Over time, the algorithm becomes gradually more accurate.

Deep learning



Deep learning

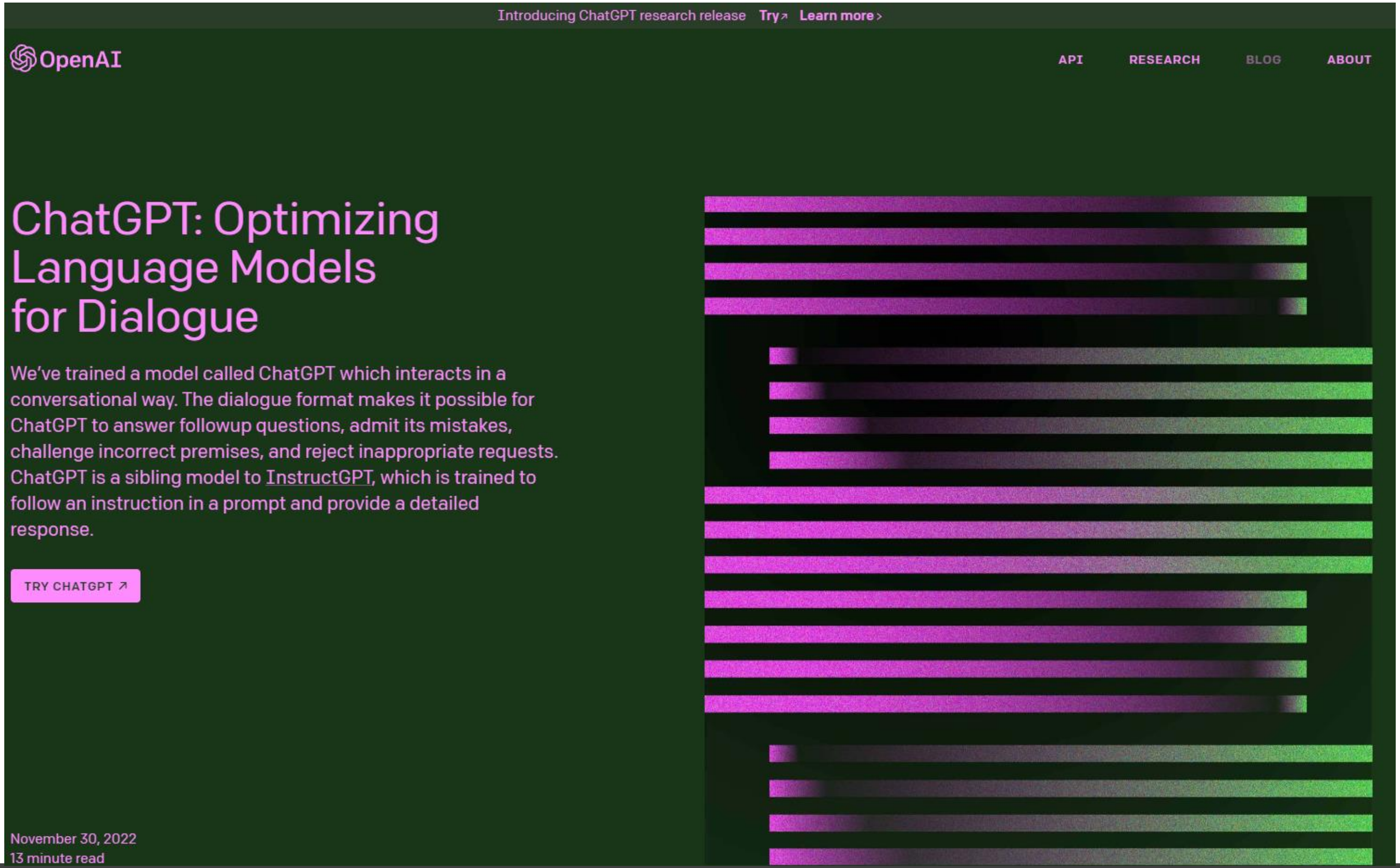


Deep learning – Application

- **Deep learning** is a subfield of machine learning that is inspired by the structure and function of the human brain, called artificial neural networks. It has been applied in a wide range of applications and industries, including:
- **Computer Vision:** Object recognition, image classification, segmentation, and generation, video analysis and understanding, etc.
- **Natural Language Processing (NLP):** Text classification, sentiment analysis, language translation, speech recognition, and generation.
- **Recommender Systems:** Personalized recommendations based on user behavior, preferences, and interaction with items.
- **Robotics:** Perception, control, and decision making tasks for autonomous robots.
- **Healthcare:** Medical image analysis, drug discovery and personalized medicine, etc.
- **Finance:** Fraud detection, risk management, algorithmic trading, etc.
- **Marketing:** Customer segmentation, targeted advertising, etc.
- **Gaming:** Game AI, decision making, and strategy development, etc.
- These are some of the most common applications of deep learning, but this field is constantly evolving and expanding to new areas.

https://chat.openai.com/chat

Deep learning – Application (ChatGPT)



https://chat.openai.com/chat

ChatGPT

How Close is ChatGPT to Human Experts? Comparison Corpus, Evaluation, and Detection

Biyang Guo[1†*], Xin Zhang[2*], Ziyuan Wang[1*], Minqi Jiang[1*], Jinran Nie[3*]
Yuxuan Ding[4], Jianwei Yue[5], Yupeng Wu[6]
[1]AI Lab, School of Information Management and Engineering
Shanghai University of Finance and Economics
[2]Institute of Computing and Intelligence, Harbin Institute of Technology (Shenzhen)
[3]School of Information Science, Beijing Language and Culture University
[4]School of Electronic Engineering, Xidian University
[5]School of Computing, Queen's University, [6]Wind Information Co., Ltd

Abstract

The introduction of ChatGPT[2] has garnered widespread attention in both academic and industrial communities. ChatGPT is able to respond effectively to a wide range of human questions, providing fluent and comprehensive answers that significantly surpass previous public chatbots in terms of security and usefulness. On one hand, people are curious about how ChatGPT is able to achieve such strength and how far it is from human experts. On the other hand, people are starting to worry about the potential negative impacts that large language models (LLMs) like ChatGPT could have on society, such as fake news, plagiarism, and social security issues. In this work, we collected tens of thousands of comparison responses from both human experts and ChatGPT, with questions ranging from open-domain, financial, medical, legal, and psychological areas. We call the collected dataset the **H**uman **C**hatGPT **C**omparison **C**orpus (**HC3**). Based on the HC3 dataset, we study the characteristics of ChatGPT's responses, the differences and gaps from human experts, and future directions for LLMs. We conducted comprehensive human evaluations and linguistic analyses of ChatGPT-generated content compared with that of humans, where many interesting results are revealed. After that, we conduct extensive experiments on how to effectively detect whether a certain text is generated by ChatGPT or humans. We build three different detection systems, explore several key factors that influence their effectiveness, and evaluate them in different scenarios. The dataset, code, and models are all publicly available at `https://github.com/Hello-SimpleAI/chatgpt-comparison-detection`.

CHATGPT GOES TO LAW SCHOOL

Jonathan H. Choi,[1] Kristin E. Hickman,[2] Amy B. Monahan,[3] Daniel Schwarcz[4]

How well can AI models write law school exams without human assistance? To find out, we used the widely publicized AI model ChatGPT to generate answers on four real exams at the University of Minnesota Law School. We then blindly graded these exams as part of our regular grading processes for each class. Over 95 multiple choice questions and 12 essay questions, ChatGPT performed on average at the level of a C+ student, achieving a low but passing grade in all four courses. After detailing these results, we discuss their implications for legal education and lawyering. We also provide example prompts and advice on how ChatGPT can assist with legal writing.

I. WHAT IS CHATGPT?

ChatGPT is an AI language model produced by OpenAI and released in late 2022.[5] GPT models, including ChatGPT, are "autoregressive," meaning that they predict the next word given a body of text. For example, given the phrase "I walked to the", a GPT model might predict that the next word is "park" with 5% probability, "store" with 4% probability, etc. The model can then repeatedly predict subsequent words (for example, "and") to compose indefinitely long bodies of text.

OpenAI has produced progressively larger language models, from GPT-1's 117 million parameters to GPT-3's 175 billion parameters.[6] One of the most important discoveries in machine learning over the past decade has been the extraordinary returns to scale when language models use more parameters and are trained on larger corpora of text.

How Does ChatGPT Perform on the Medical Licensing Exams? The Implications of Large Language Models for Medical Education and Knowledge Assessment

Aidan Gilson[1,2], Conrad Safranek[1], Thomas Huang[2], Vimig Socrates[1,3], Ling Chi[1], R. Andrew Taylor[1,2], and David Chartash[1,4]

[1]Section for Biomedical Informatics and Data Science, Yale University School of Medicine
[2]Department of Emergency Medicine, Yale School of Medicine
[3]Program of Computational Biology and Bioinformatics, Yale University
[4]School of Medicine, University College Dublin - National University of Ireland, Dublin

ABSTRACT

Background: ChatGPT is a 175 billion parameter natural language processing model which can generate conversation style responses to user input.
Objective: To evaluate the performance of ChatGPT on questions within the scope of United States Medical Licensing Examination (USMLE) Step 1 and Step 2 exams, as well as analyze responses for user interpretability.
Methods: We used two novel sets of multiple choice questions to evaluate ChatGPT's performance, each with questions pertaining to Step 1 and Step 2. The first was derived from AMBOSS, a commonly used question bank for medical students, which also provides statistics on question difficulty and the performance on an exam relative to the userbase. The second, was the National Board of Medical Examiners (NBME) Free 120-question exams. After prompting ChatGPT with each question, ChatGPT's selected answer was recorded, and the text output evaluated across three qualitative metrics: logical justification of the answer selected, presence of information internal to the question, and presence of information external to the question.
Results: On the four datasets, AMBOSS-Step1, AMBOSS-Step2, NBME-Free-Step1, and NBME-Free-Step2, ChatGPT achieved accuracies of 44%, 42%, 64.4%, and 57.8%. The model demonstrated a significant decrease in performance as question difficulty increased (P=.012) within the AMBOSS-Step1 dataset. We found logical justification for ChatGPT's answer selection was present in 100% of outputs. Internal information to the question was present in >90% of all questions. The presence of information external to the question was respectively 54.5% and 27% lower for incorrect relative to correct answers on the NBME-Free-Step1 and NBME-Free-Step2 datasets (P<=.001).
Conclusion: ChatGPT marks a significant improvement in natural language processing models on the tasks of medical question answering. By performing at greater than 60% threshold on the NBME-Free-Step-1 dataset we show that the model is comparable to a third year medical student. Additionally, due to the dialogic nature of the response to questions, we demonstrate ChatGPT's ability to provide reasoning and informational context across the majority of answers. These facts taken together make a compelling case for the potential applications of ChatGPT as a medical education tool.

ChatGPT

THE LANCET

Comment

Generating scholarly content with ChatGPT: ethical challenges for medical publishing

The impact of generative artificial intelligence (AI) on medical publishing practices is currently unknown. However, as our experiences underline, generative AI could have substantial ethical implications.

ChatGPT (OpenAI, San Francisco, CA, USA) is an AI chatbot released in November, 2022.[1] Developed using human feedback and freely accessible, the platform has already attracted millions of interactions.[2] When presented with a query, ChatGPT will automatically generate a response, which is based on thousands of internet sources, often without further input from the user. Resultantly, individuals have reportedly used ChatGPT to formulate university essays and scholarly articles[3] and, if prompted, the system can deliver accompanying references. Given these accounts and its popular usage, we requested that ChatGPT write a Comment for *The Lancet Digital Health* about AI and medical publishing ethics. We then asked ChatGPT how the editorial team should address academic content produced by AI. The results make for interesting reading (appendix).

The functionality of ChatGPT highlights the growing necessity of implementing robust AI author guidelines in scholarly publishing. Ethical considerations abound

journals, alongside other major publishers, have stated that AI cannot be listed as an author and its use must be properly acknowledged.[7]

ChatGPT is available to use without cost.[1] However, OpenAI's leadership have affirmed that free use is temporary and the product will eventually be monetised.[8] One commercial option for the platform could conceivably involve some form of paywall, which might entrench existing international inequalities in scholarly publishing. Although institutions in socioeconomically advantaged areas could probably afford access, those in low-income and middle-income countries might not be able to, thus widening existing disparities in knowledge dissemination and scholarly publishing.

In our opinion, as the program remains freely available in the short term, ChatGPT's ease of use and accessibility could substantially increase scholarly output. ChatGPT might democratise the dissemination of knowledge since the chatbot can receive and produce copy in multiple languages, circumventing English-language requirements that can be a publishing barrier for speakers of other languages. Nonetheless, the functionality of ChatGPT has the capacity to cause harm by producing misleading or inaccurate content,[3] thereby

Lancet Digit Health 2023

Published Online February 6, 2023 https://doi.org/10.1016/S2589-7500(23)00019-5

See Online for appendix

Reviews and Commentary
Editorial

Radiology

ChatGPT Is Shaping the Future of Medical Writing but Still Requires Human Judgment

Paper Link

Felipe C. Kitamura

Author Affiliations

Published Online: Feb 2 2023 | https://doi.org/10.1148/radiol.230171

Free Access

Reviews and Commentary
Editorial

ChatGPT and Other Large Language Models Are Double-edged Swords

Paper Link

Yiqiu Shen[1], Laura Heacock[2], Jonathan Elias[3], Keith D. Hentel[4], Beatriu Reig[2], George Shih[4], Linda Moy[2]

Author Affiliations

Published Online: Jan 26 2023 | https://doi.org/10.1148/radiol.230163

Free Access

Reviews and Commentary
Perspectives

ChatGPT and the Future of Medical Writing

Paper Link

Som Biswas

Author Affiliations

Published Online: Feb 2 2023 | https://doi.org/10.1148/radiol.223312

Reviews and Commentary
Editorial

Free Access

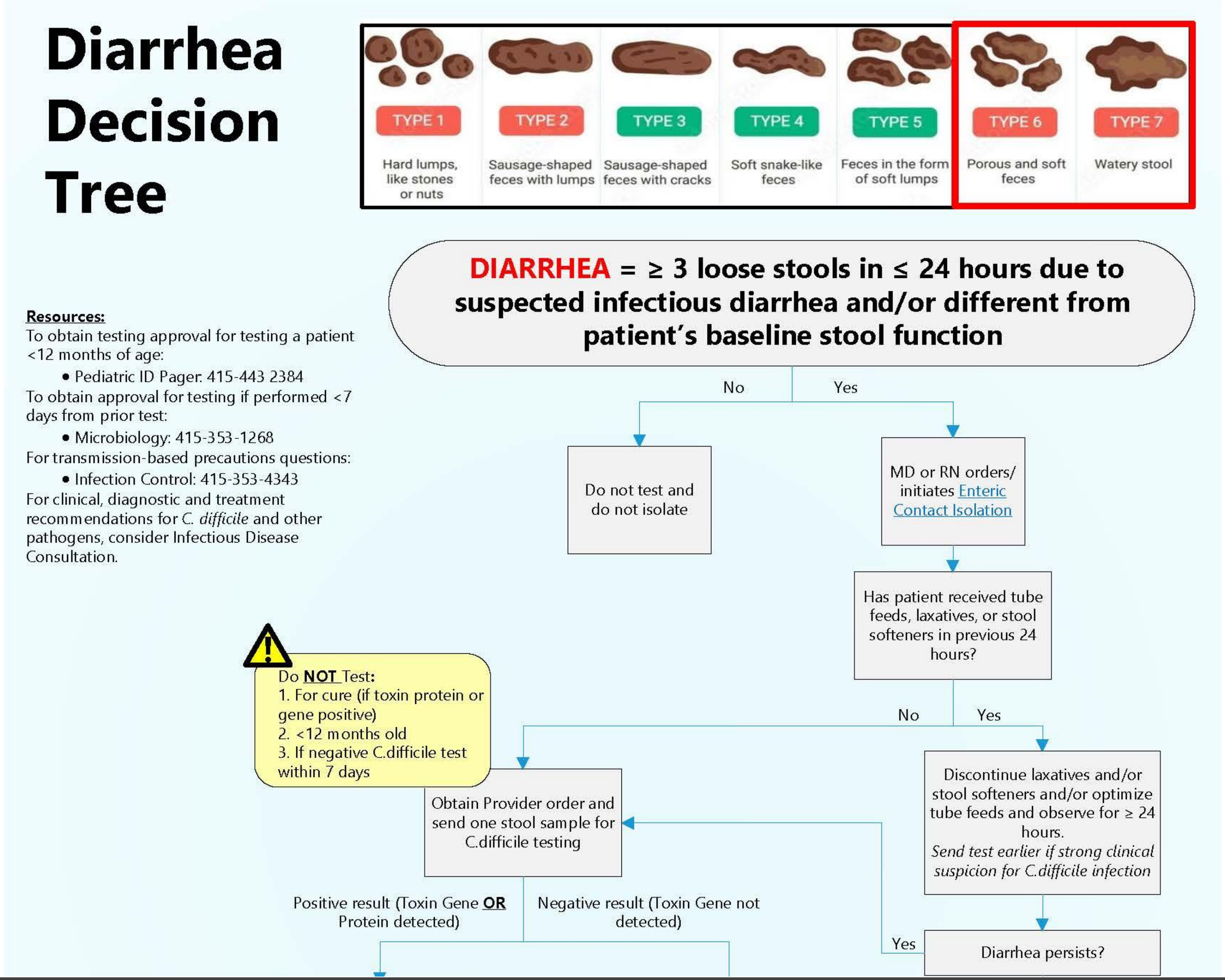
ChatGPT Is Shaping the Future of Medical Writing but Still Requires Human Judgment

Felipe C. Kitamura

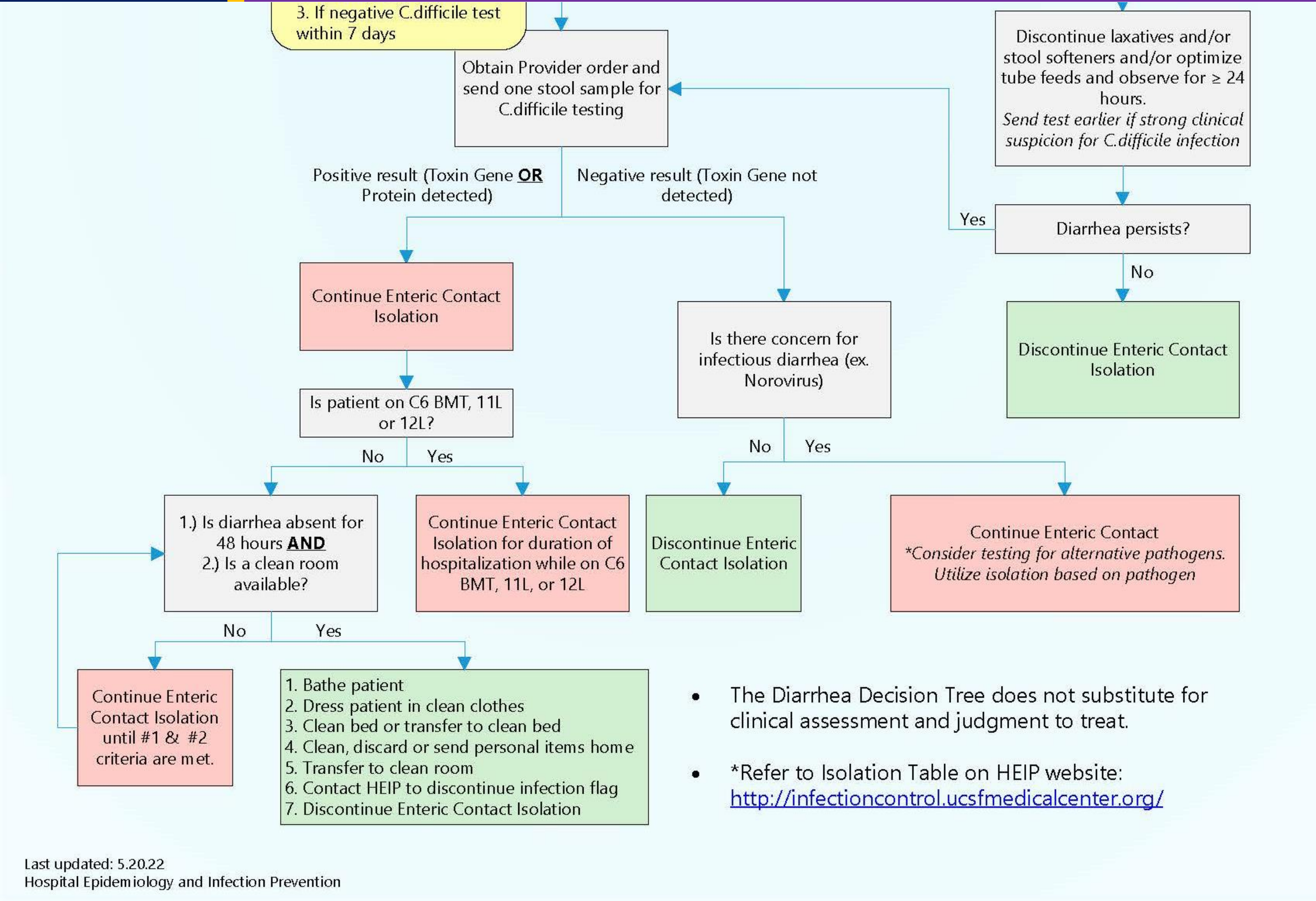
Author Affiliations

Published Online: Feb 2 2023 | https://doi.org/10.1148/radiol.230171

- **Medical writing:** scientific articles, radiology reports, regulatory documents, and patient educational materials.
- The training process of GPT involves asking the model to **predict the next element in a sentence**.
- That enables GPT to generate random free text, mimicking the training text (and being biased by it).
- ChatGPT is a more complex model that requires two more steps.
- After pre-training GPT with a large corpus, GPT is trained to predict the answer to a given question.



https://infectioncontrol.ucsfmedicalcenter.org/diarrhea-decision-tree



https://infectioncontrol.ucsfmedicalcenter.org/diarrhea-decision-tree

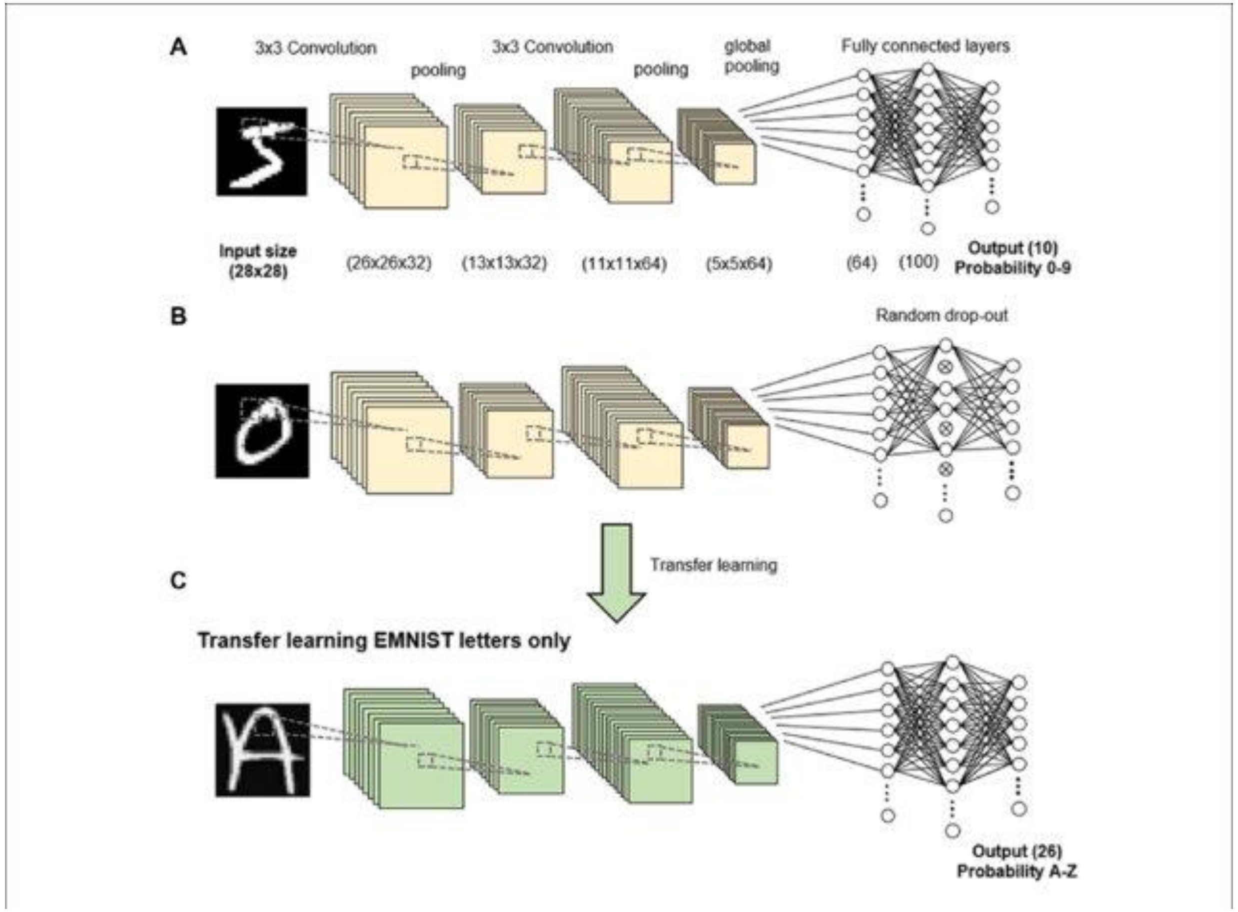
Som Biswas (2023) ChatGPT and the Future of Medical Writing. *Radiology*. https://doi.org/10.1148/radiol.223312

- **Ethics:** The use of AI in writing raises concerns about authorship and accountability for the content that is generated. Although chat GPT generates articles that have less plagiarism, they are not completely free of it and need editing by human authors. Also letters of recommendation and personal statements based on works created by chat GPT can be raise authenticity concerns.
- **Legal issues:** There are several legal issues to consider when using ChatGPT or other AI-powered language models:
  - ➢ **Copyright:** When AI-generated text is used for commercial purposes, it's important to ensure that the use of AI-generated text does not infringe on any existing copyrights.
  - ➢ **Compliance:** In some fields such as healthcare and legal, the use of AI-generated text may be subject to regulations and compliance. Currently no laws exist regarding utilization of AI in medical literature.
  - ➢ **Medico-legal issues**; Provider documentation into the patient's medical record including radiology reports created by AI could lead to errors and thus lawsuits. Questions of accountability will then arise regarding these reports.

Som Biswas (2023) ChatGPT and the Future of Medical Writing. *Radiology*. https://doi.org/10.1148/radiol.223312

- **Innovation:** As chat GPT is based on prior data fed to it, eventually it will lead to repetitive text generation and lack of creativity. ChatGPT and other AI-generated text may lack the creativity and originality that human authors bring to their work. Also easy automated text generation can also lead to less engaged students regarding coursework and assignments, in medical schools and colleges across the world.
- **Accuracy:** There is a concern that AI-generated text may not be accurate or may contain errors. The current version of ChatGPT does not offer any assessment of content accuracy.
- **Bias:** AI models are trained on large amounts of data, which may include bias. Therefore, there is a concern that the text generated by AI may perpetuate or amplify bias.
- **Transparency:** The use of AI in the writing process and identification of text that has been generated by a machine should be made clear.

Number/ plate detection

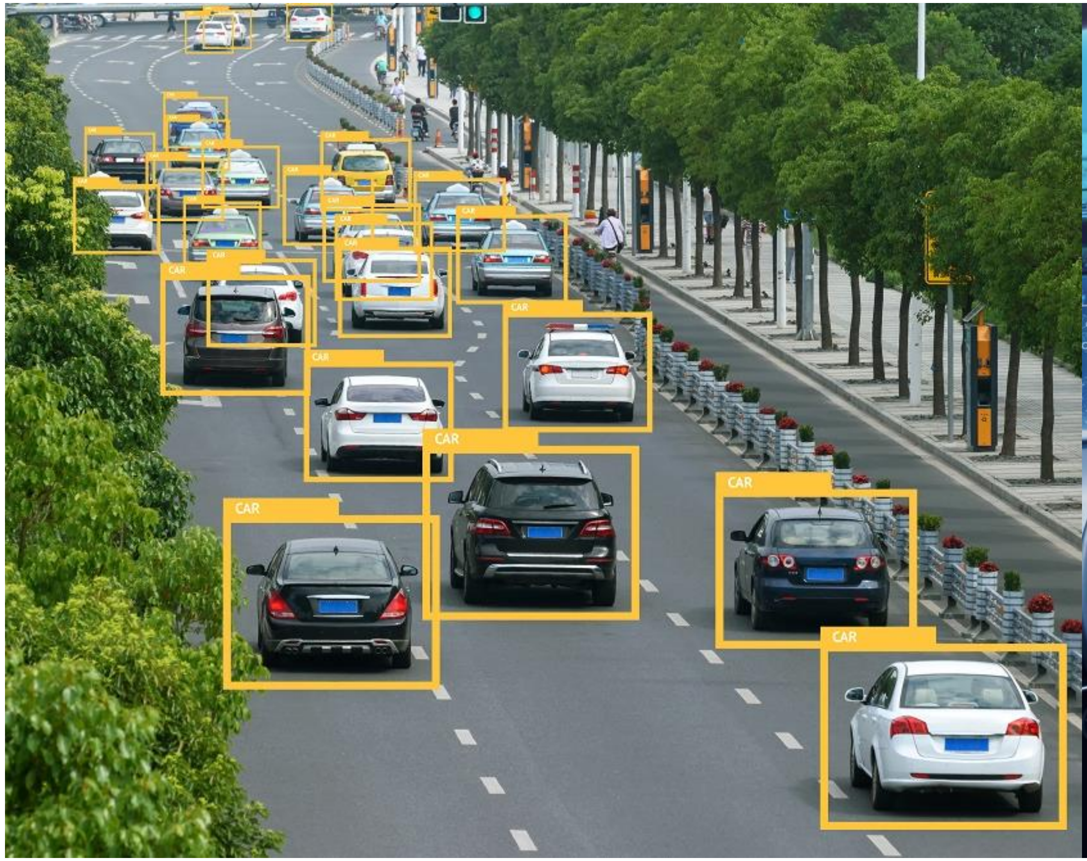


Standard writing detection ➔
Hand writing detection ➔
ETC

https://news.tvbs.com.tw/local/1130681

https://www.researchgate.net/publication/343547826_Dropout_in_Neural_Networks_Simulates_the_Paradoxical_Effects_of_Deep_Brain_Stimulation_on_Memory/figures?lo=1

Object Detection



https://www.acrosser.com.tw/zh-tw/Solutions/Autonomous-Driving-Servers/
https://tw.cyberlink.com/faceme/insights/articles/478/automatic-license-plate-recognition-facial-recognition

Stock price prediction

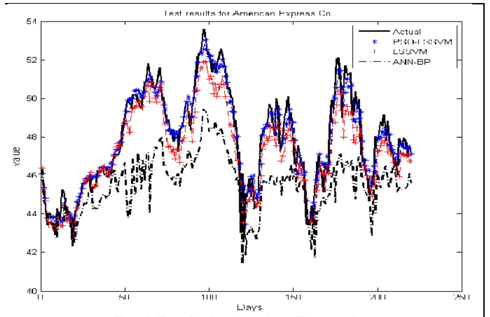


Fig. 6: Results for American Express Co

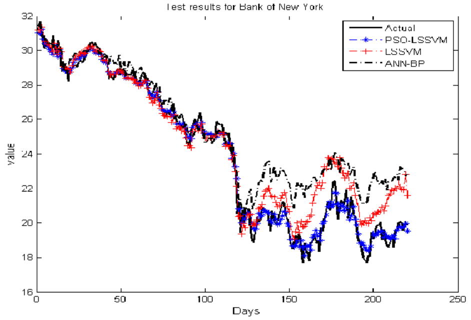


Fig. 7: Results for Bank of New York

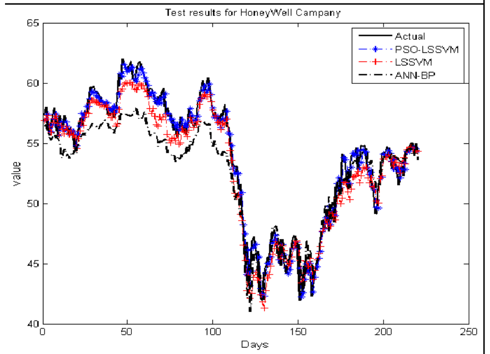


Fig. 8: Results for Honeywell company

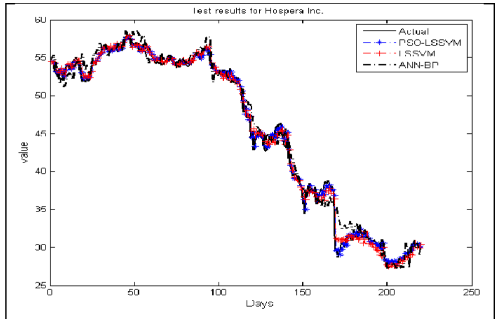


Fig. 9: Results for Hospera Company

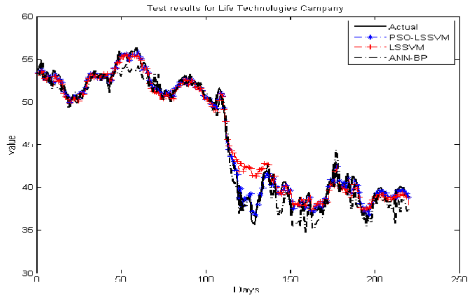


Fig. 10: Results for Life Technologies company

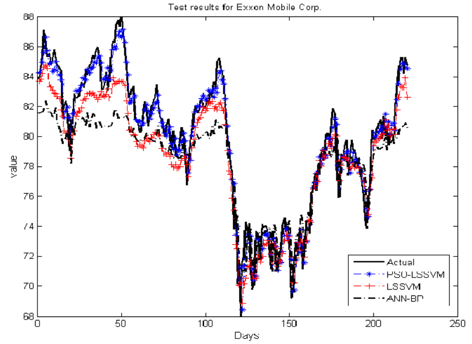


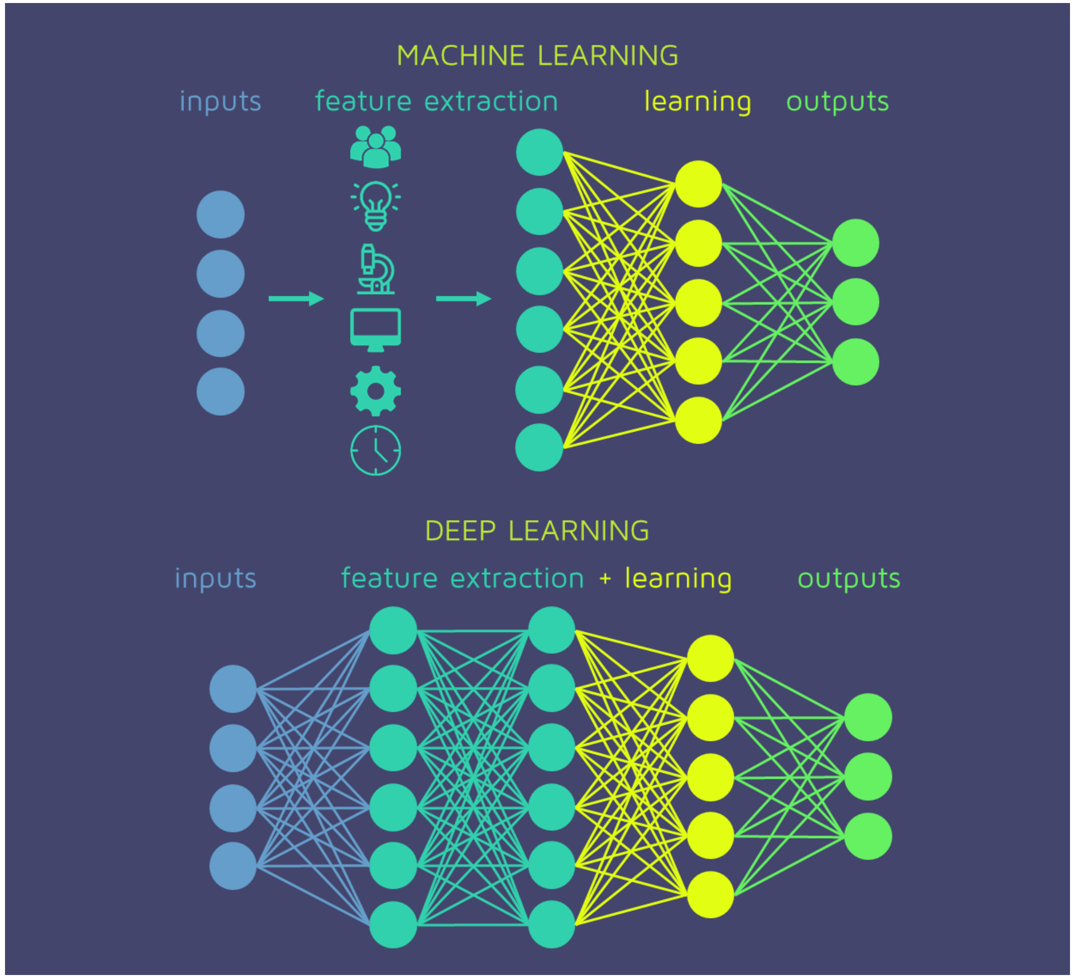
Fig. 11: Results for Exxon Mobile company

https://www.semanticscholar.org/paper/A-Machine-Learning-Model-for-Stock-Market-Hegazy-Soliman/ab3297f44f6cdaa778b97efe930f6534b3edc92f

Machine learning & deep learning

- If deep learning is a subset of machine learning, how do they differ? Deep learning distinguishes itself from classical machine learning by the type of data that it works with and the methods in which it learns.
- Machine learning algorithms leverage structured, labeled data to make predictions—meaning that specific features are defined from the input data for the model and organized into tables. This doesn’t necessarily mean that it doesn’t use unstructured data; it just means that if it does, it generally goes through some pre-processing to organize it into a structured format.

Machine learning & deep learning

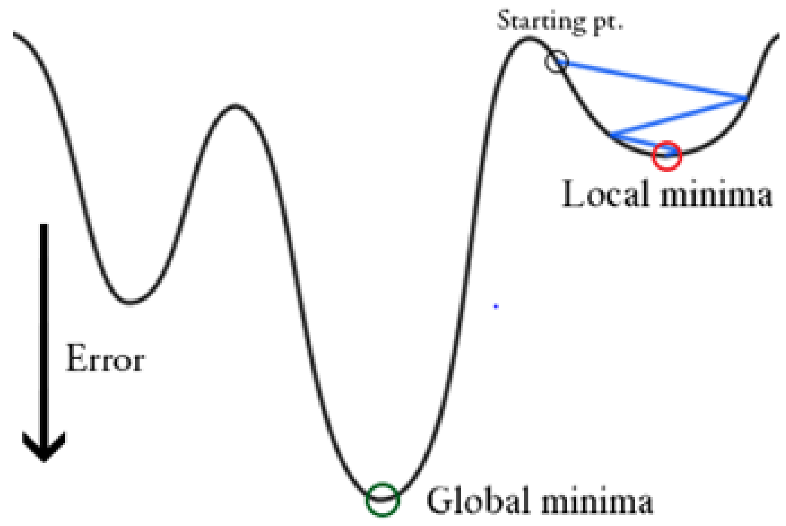


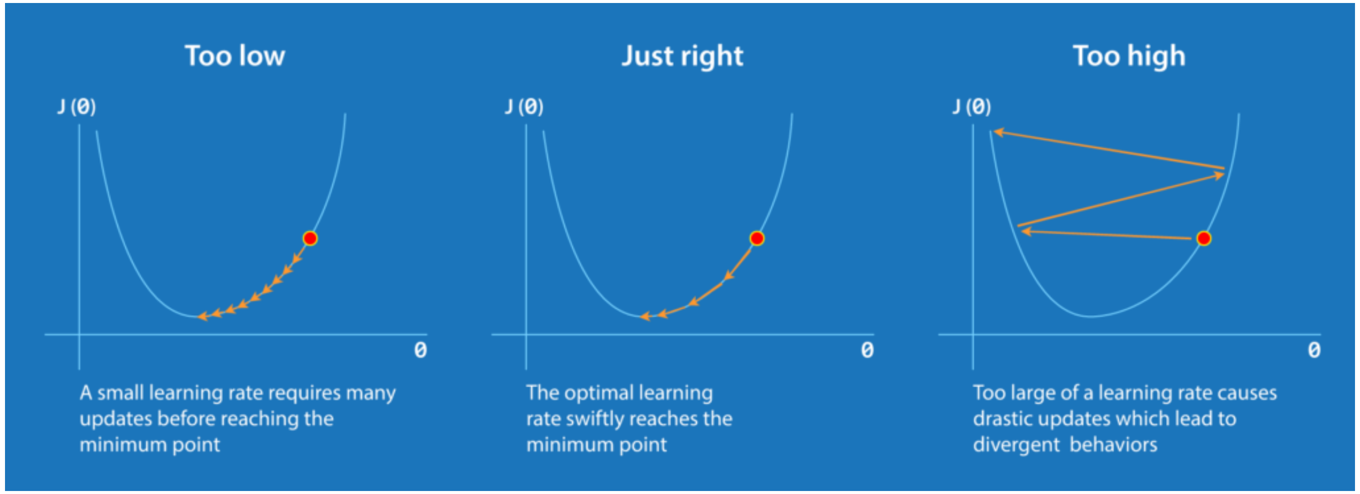
Apparently, the operations that Neural Networks perform during the training phase are equivalent to "searching for meaningful combinations of variables that produce better performance results on the training data". I have heard that this is in some way similar to Principal Component Analysis (PCA) and Kernel Methods, seeing as these methods combine many different existing features into new features that have "more meaningful representation".

Deep learning solves this central problem in representation learning by introducing representations that are expressed in terms of other, simpler representations. Deep learning enables the computer to build complex concepts out of simpler concepts.

https://stats.stackexchange.com/questions/562466/neural-networks-automatically-do-feature-engineering-how

Optimization





https://www.mygreatlearning.com/blog/gradient-descent/#types-of-gradient-descent
https://www.fromthegenesis.com/gradient-descent-part1/

Potential issues

- **Overfitting:** Deep learning models have the tendency to memorize the training data, leading to overfitting and poor generalization performance on new, unseen data. This can be mitigated using techniques such as early stopping, regularization, or dropout.
- **Computational Complexity:** Training deep learning models can be computationally expensive, requiring large amounts of data and computational resources. This can make it challenging to scale up deep learning models to handle large amounts of data, or to deploy them on resource-constrained devices.
- **Data Bias:** Deep learning models can learn and amplify existing biases in the training data, leading to biased predictions or decisions. This can be a particularly important issue in applications such as facial recognition, where the model may be trained on data that is biased towards a particular race or gender.

https://chat.openai.com/chat

Potential issues

- **Explanation and Interpretability:** Deep learning models can be difficult to understand and interpret, making it challenging to determine why the model is making a particular prediction or decision. This can be an issue in applications where transparency and accountability are important, such as in medical diagnosis or criminal justice.
- **Adversarial Examples:** Deep learning models can be vulnerable to adversarial examples, where small, carefully crafted changes to the input data can cause the model to make incorrect predictions. This is a security concern in applications such as image recognition, where an attacker could potentially manipulate the input to cause the model to misbehave.

https://chat.openai.com/chat

Paper Reading

https://arxiv.org/ftp/arxiv/papers/2302/2302.02083.pdf

Theory of Mind May Have Spontaneously Emerged in Large Language Models

Authors: Michal Kosinski*[1]

Affiliations:

[1]Stanford University, Stanford, CA94305, USA

*Correspondence to: michalk@stanford.edu

Abstract: Theory of mind (ToM), or the ability to impute unobservable mental states to others, is central to human social interactions, communication, empathy, self-consciousness, and morality. We administer classic false-belief tasks, widely used to test ToM in humans, to several language models, without any examples or pre-training. Our results show that models published before 2022 show virtually no ability to solve ToM tasks. Yet, the January 2022 version of GPT-3 (davinci-002) solved 70% of ToM tasks, a performance comparable with that of seven-year-old children. Moreover, its November 2022 version (davinci-003), solved 93% of ToM tasks, a performance comparable with that of nine-year-old children. These findings suggest that ToM-like ability (thus far considered to be uniquely human) may have spontaneously emerged as a byproduct of language models' improving language skills.

Assignment

- Given at least three examples for machine learning and deep learning with appropriate reasons.

Examples	Machine learning	Deep learning
Ex 1		
Ex 2		
Ex 3		

Thank you for your attention!