



# CloudAEye Webinar series

---

How to become a data-driven  
enterprise with AI

(episode 1)

# Introduction: Who are we

CloudAEye

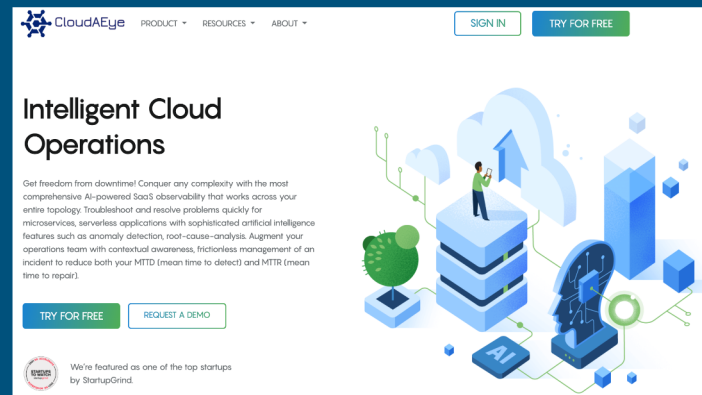
AI-driven observability for cloud-native applications


Presenter Bio

Co-Founder and CTO of CloudAEye

AI and recommender systems experience

Led the AI-academy at LinkedIn





**Badrul M. Sarwar**  
Machine Learning Scientist, [LinkedIn](#)  
Verified email at linkedin.com - [Homepage](#)  
[Recommender systems](#) [collaborative filtering](#) [machine learning](#) [deep learning](#) [NLP](#)

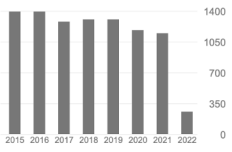
[FOLLOW](#)

TITLE	CITED BY	YEAR
<input type="checkbox"/> <a href="#">Item-based collaborative filtering recommendation algorithms</a> B Sarwar, G Karypis, J Konstan, J Riedl Proceedings of the 10th international conference on World Wide Web, 285-295	10935	2001
<input type="checkbox"/> <a href="#">Analysis of recommendation algorithms for e-commerce</a> B Sarwar, G Karypis, J Konstan, J Riedl Proceedings of the 2nd ACM Conference on Electronic Commerce, 158-167	2856	2000
<input type="checkbox"/> <a href="#">Application of dimensionality reduction in recommender system-a case study</a> B Sarwar, G Karypis, J Konstan, J Riedl ACM WebKDD 2000 Workshop	2213	2000

**Cited by**

[VIEW ALL](#)

	All	Since 2017
Citations	20037	6488
h-index	19	13
i10-index	22	13



Bar chart showing citation counts from 2015 to 2022. The y-axis ranges from 0 to 1400. The chart shows a general downward trend in citations over the years, with a significant drop in 2022.

# Agenda

---

- Organization of the webinar series

  - What to expect, topics covered, tools and contents provided

- Data-driven Enterprise

  - Data-driven organizations and the role of AI

- Introduction to AI concepts

  - AI/ML/DL/DS

  - Introduction to basics of ML: supervised and unsupervised learning

  - Supervised and unsupervised learning basics: Regression, classification and Clustering etc.

- Code examples

- Summary and QA

# Organization of the Webinar series

---

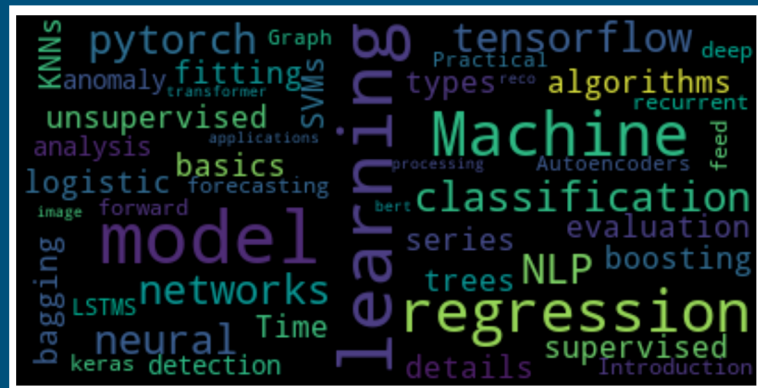
Cadence:	Bi-weekly meeting of 30 minutes
Duration:	Approximately 15-20 episodes
Focus:	Basic concepts of AI/ML and how organizations can utilize them
Format:	Interactive, presentation, code examples, Q/A, assignment
Resources:	Github and colaboratory for code, slack-channel for discussion

# Webinar series: Tentative Topics

- Introductory concepts
- Machine learning models: types of models, model fitting basics, supervised and unsupervised learning, regression and classification, evaluation
- Classification algorithms details: logistic regression, trees: bagging/boosting, SVMs, KNNs, Bayesian
- Time-series analysis, anomaly detection, forecasting

## Practical and hands-on

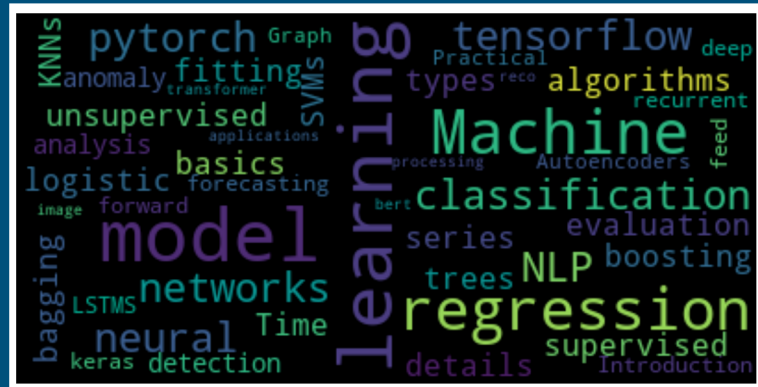
### Example code and assignments



# Webinar series: Tentative Topics (contd.)

- Introduction to deep learning: feed forward and recurrent neural networks, LSTMS, Autoencoders, Graph neural networks
- Personalization/recommendation: Matrix factorization, wide and deep models, GNN
- Practical applications: image processing, NLP, recommender systems, AIOps
- Transfer learning, self-supervised learning
- Emerging concepts: privacy, fairness, explainability

State-of-the-art models and techniques  
Practical large-scale examples



# Data-driven Enterprises

---

## Data-driven culture

- We are submerged in data (Big data)
- Data can be leveraged to make better decisions, gauge and improve customer sentiment, streamline operations etc.

## What are the traits of a data-driven enterprise?

- Fosters a data-driven culture starting from the top
- Metrics are carefully chosen
- Employees are encouraged to be data-centric and trained on latest tech
- Provides a robust infrastructure for collecting, analyzing and serving

For more info: <https://hbr.org/2020/02/10-steps-to-creating-a-data-driven-culture>

# Data-driven Enterprises: Challenges

---

Building a Data-driven culture is hard

- Cultural shift is harder than you think
- It's not about the technology it's about the people

Barriers of becoming a data-driven enterprise

- Shortage of talent—leadership to worker
- Sheer volume of data and lack of infrastructure
- Lack of clean usable data
- Concerns about privacy

For more info: <https://hbr.org/2022/02/why-becoming-a-data-driven-organization-is-so-hard>



# How to overcome these challenges

---

There is no short term solution, it takes time

- Think different

There is no shortage of analytic algorithms. These need to be matched by critical thinking, human judgement, and a view to creative innovation.

- Fail fast, learn faster

Companies that are prepared for faster iterative learning — fail fast, learn faster — will gain insight and knowledge before their competitors.

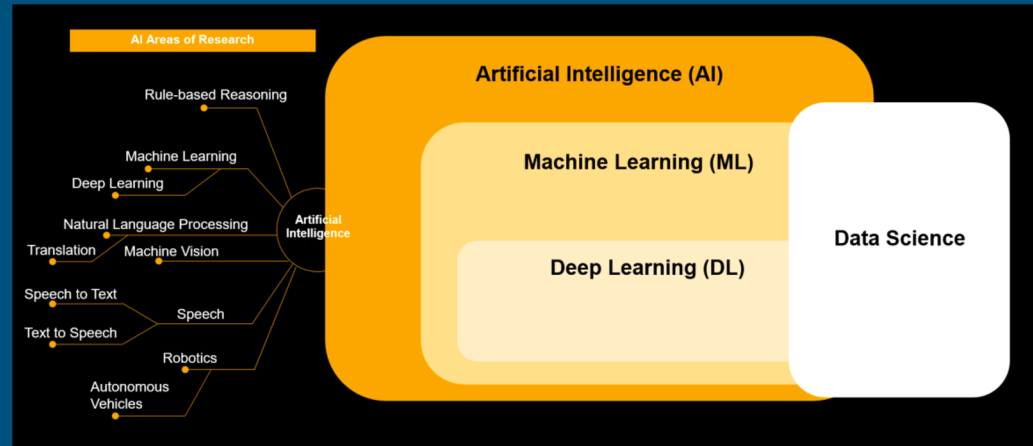
- Focus on the long-term

Becoming data-driven is a process. It takes time

# AI, ML, DL, DS Demystified

AI: Intelligence exhibited by machines. Broadly defined to include any simulation of human intelligence

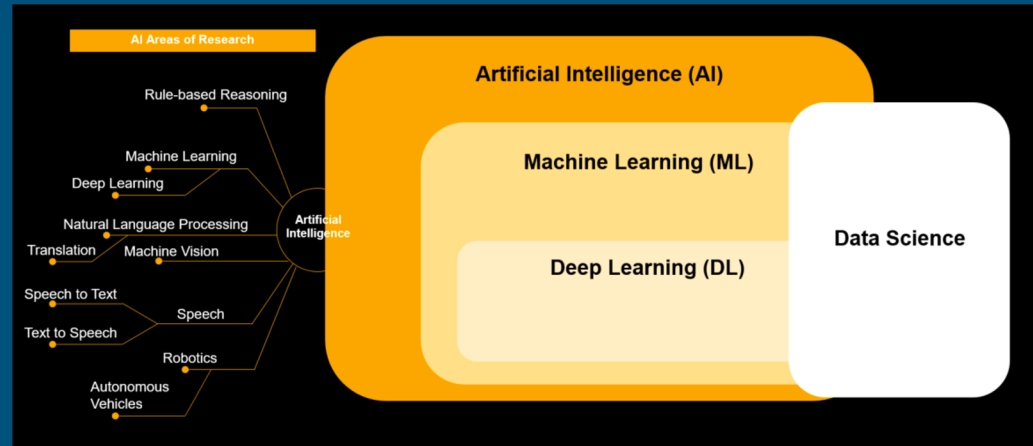
ML: Subfield of AI that aims to teach computers the ability to do tasks with data, without explicit programming



# AI, ML, DL, DS Demystified

DL: Subfield of ML that uses specialized techniques involving multi-layer (2+) artificial neural networks

DS: Scientific methods, algorithms and systems to extract knowledge or insights from big data



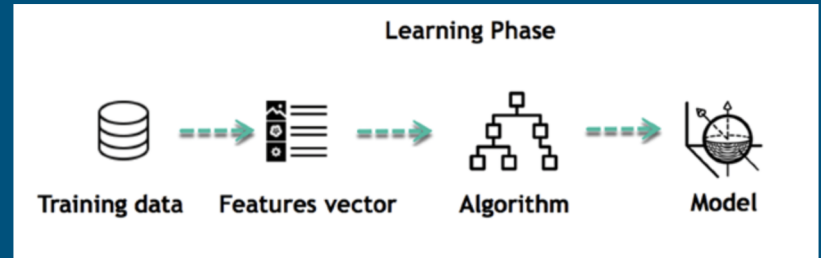
For this webinar series we will focus mostly on ML/DL algorithms

# Machine learning algorithms

Machine learning algorithms aim to build a model or hypothesis based on the input data

Depending on the type on input ML algorithms can be

Supervised or unsupervised



# Supervised ML algorithms

When the data is “labeled” or “tagged” for the correct outcome or value:

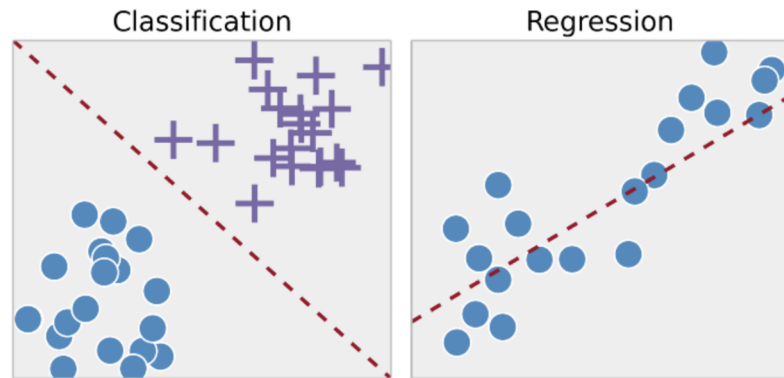
Input looks like  $(X, y)$

Where  $X$  is the set of input features and  $y$  is the outcome variable that is provided

Based on  $y$  we have two classes of supervised learning:

1. Regression, where  $y$  is continuous ex: linear regression, tree-based regression, SVR
2. Classification, where  $y$  is discrete. ex: logistic regression, decision trees, SVM, KNN, deep networks

Types of Supervised Machine Learning Techniques



# Unsupervised ML algorithms

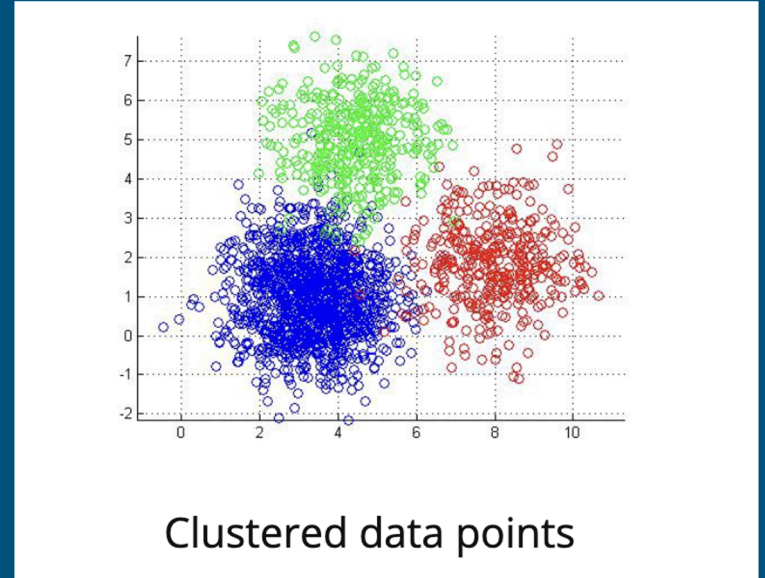
When the data is NOT “labeled” or “tagged”

Input looks like (X)

Where X is the set of input features

Types of unsupervised learning:

Clustering, density estimation,  
dimensionality reduction, representation  
learning, auto-encoders



# Other types of ML

---

## 1. Semi-supervised learning

Some labeled data and mostly unlabeled data are provided, the algorithm uses the labeled data to reason about the unlabeled data

## 1. Self-supervised learning

Common-sense learning using lots of unsupervised data e.g., text, images

## 1. Reinforcement learning

Learning by trials and errors, given a state, a set of actions and reward function these algorithms learn to achieve intelligence through exploration

# Coding example

---

Link to the colab Notebook

<https://colab.research.google.com/drive/15xzHZU9G2eiJlyhabFqPFIvAAWki68l1?usp=sharing>



# Thank You

---

You have DevOps pain points?

We like to hear from you

Looking for exciting opportunities?

We are hiring

scientists      AI/ML engineers and applied

Cloud Engineers

UI/UX designers



**CloudAEye**