Open in app ↗

Medium   🔍 Search   🔔   Ⓢ

✦   Get unlimited access to the best of Medium for less than $1/week.   **Become a member**   ✕

Introduction to Vectors for Data Science

Vishal Kumar · Follow
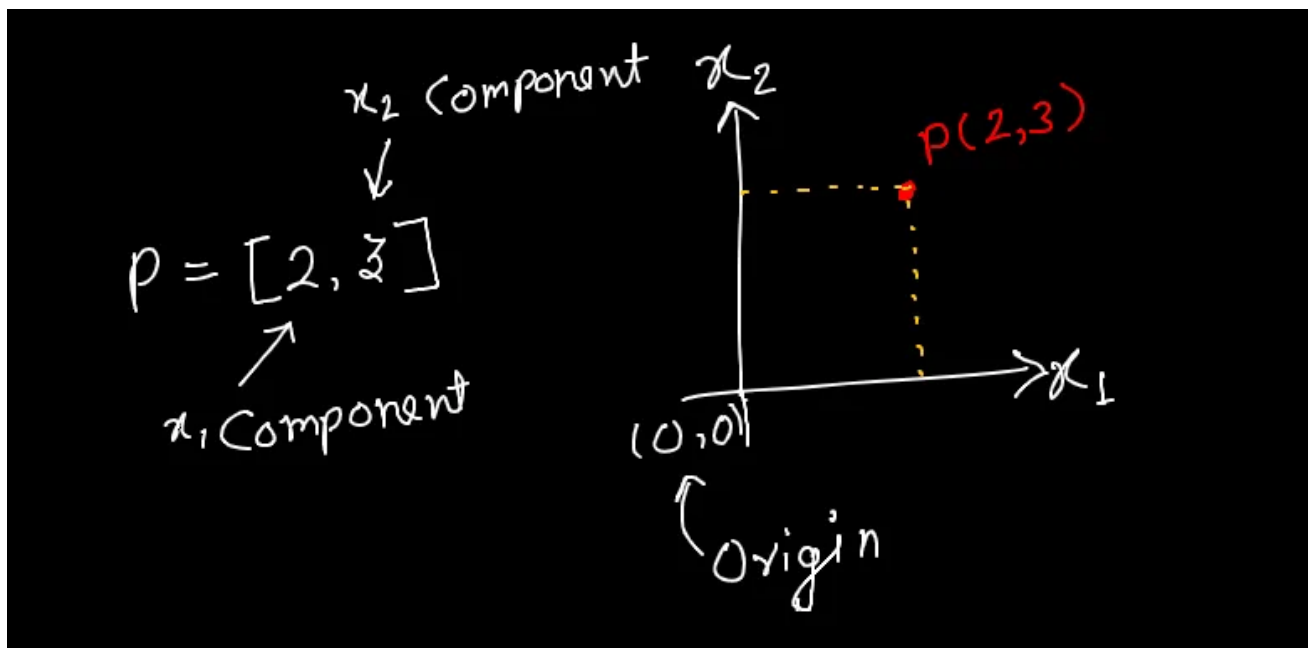5 min read · Apr 10, 2022
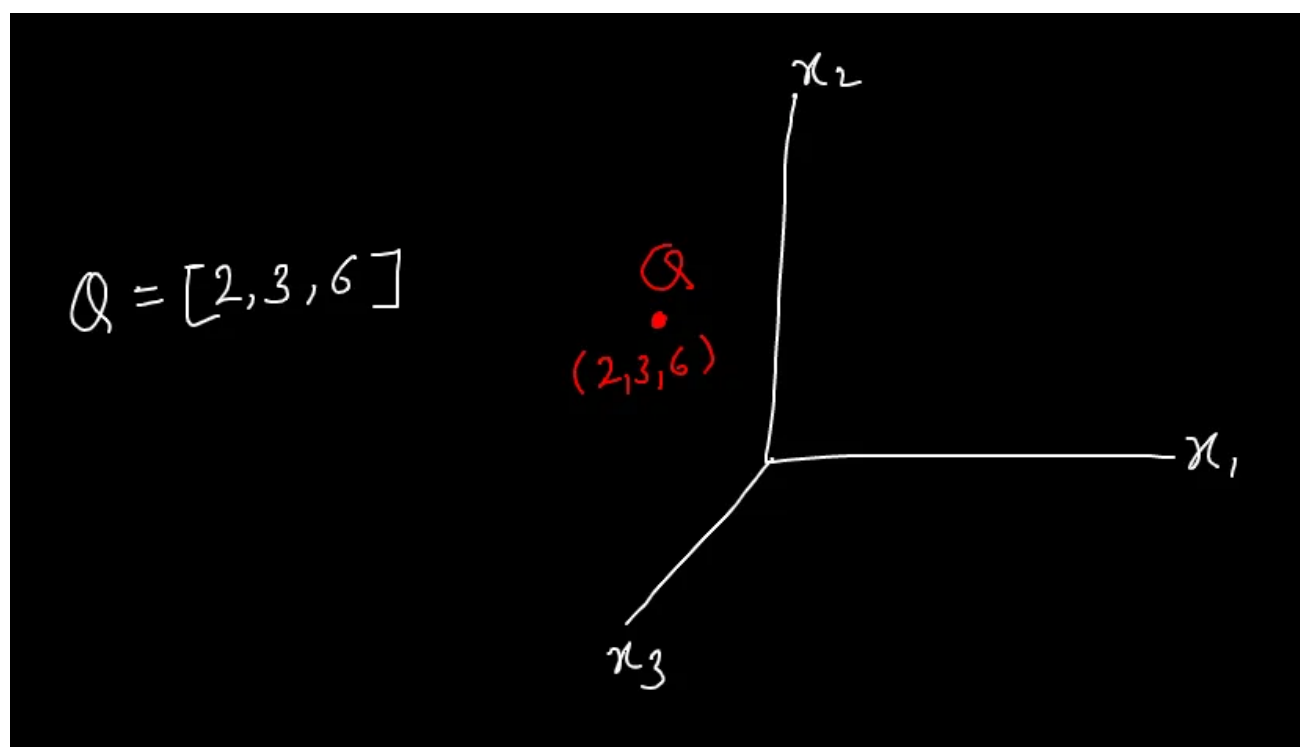
▶ Listen   ↥ Share   ••• More

Vectors in Data Science tell the properties of a data point in different dimensions. Different components of a data point forms a vector each component is related to one dimension.

Point/Vector P in 2-D space

In the above image point **P** is a vector in 2 Dimension space with x1 and x2 component or (x,y component).

(0,0) represents the origin.

Point/Vector Q in 3-D space

In the above image point Q is a vector in 3-D space with x1,x2,x3 components. A mosquito at one position in a square room is like a data point in 3-D space 😄.

Similarly we can have N-Diminsion in a vector but it's hard to plot N-D vector on a 2D surface. A vector in N-D will look like this V = [x1,x2,x3,......,xN]

Distance of a point from origin:

Let's see how we can calculate distance of a point from origin in a space.
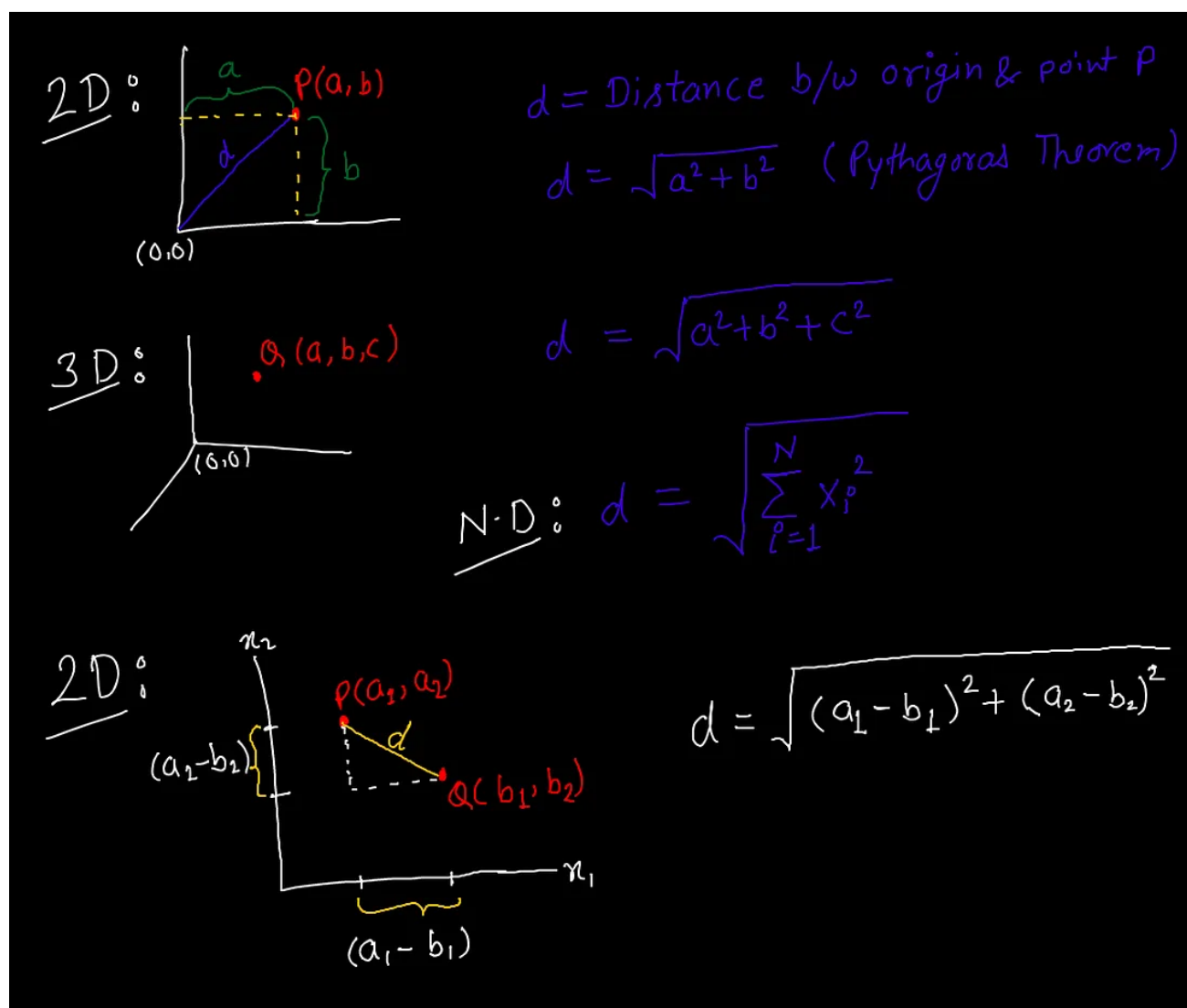
Distance between two points in 2D space

In the above image, we are calculating the distance "d" between two points "P" and "Q". Calculation of distance between two points in a space is similar to calculating the distance of a point from origin. You can consider in the above example if let, point Q is origin then the co-ordinates of point Q would be (0,0) and the same formula will get converted to the previous formula we used to find distance of point from origin.

$$3D: \quad P(a_1, a_2, a_3)$$
$$Q(b_1, b_2, a_3)$$
$$d = \sqrt{(a_1-b_1)^2 + (a_2-b_2)^2 + (a_3-b_3)^2}$$

$$N\text{-}D: \quad P(a_1, a_2, a_3, \ldots a_n)$$
$$Q(b_1, b_2, b_3, \ldots b_n)$$
$$d = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2}$$

A row vector has one row and n columns. Distance between two points in 3D and N-D space

$$\text{Row Vector:-}$$
$$A = [a_1, a_2, a_3, \ldots, a_n]_{(1 \times n)}$$

— Columns

rows

Row Vector Representation

Column Vector:

A column vector has one column and n rows.

$$\text{Column Vector:}$$

$$B = \begin{bmatrix} b_1, \\ b_2, \\ b_3, \\ \vdots \end{bmatrix} \quad \text{— Columns}$$

$$A = [a_1, a_2, a_3, \ldots a_n]$$

$$B = [b_1, b_2, b_3, \ldots b_n]$$

$$A + B = [a_1 + b_1, a_2 + b_2, a_3 + b_3, \ldots a_n + b_n]$$

Addition of two vectors A & B

In the above image we can see how to add two vectors.

Multiplication of two vector:

There are two type of multiplication we can perform on vectors, Dot Product and Cross Product, For Data Science related study Cross Product is not used frequently so we will focus on Dot Product.

Transpose:

Before performing the dot product on two vectors perform transpose operation on one of the vector iff both the vectors are of same representation e.g. both the vectors are row vectors. Transpose of a vector converts the row vector to column vector and column vector to row vector.

Transpose of vector A

In the above image vector A^T is transpose of vector A.

Dot Product:

We represent the dot product of two vectors by putting a dot between the vectors e.g. (A . B).

Note: For performing the dot product between two vectors number of column in vector 1 and number of row in vector 2 should be same. Which means dimension of both the vectors should be same. Before performing the dot product perform the transpose operation on one of the vector iff both the vectors are of same representation e.g. both the vectors are row vectors.

Geometric Intuition Behind Dot Product:

$$\text{Dot Product :-} \qquad A \cdot B$$

$$\longleftarrow \text{Dot}$$

$$A = [a_1, a_2, a_3, \ldots \ldots a_n]_{(1 \times N)} \qquad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}_{(N \times 1)}$$
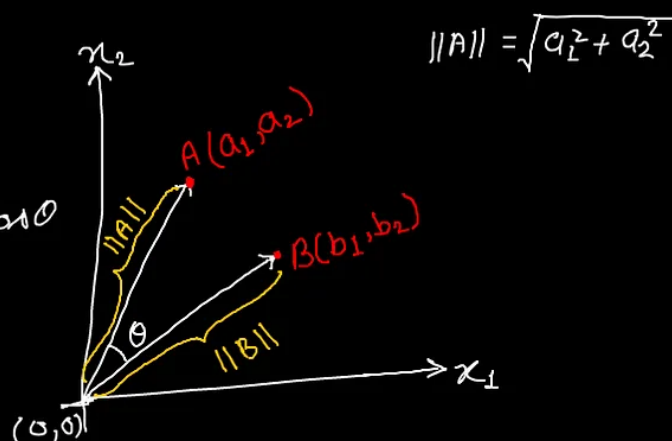
$$A \cdot B = a_1 b_1 + a_2 b_2 + a_3 b_3 + \ldots \ldots + a_n b_n$$

Dot Product of vector A and B

$$A \cdot B = \|A\| \|B\| \cos \theta$$

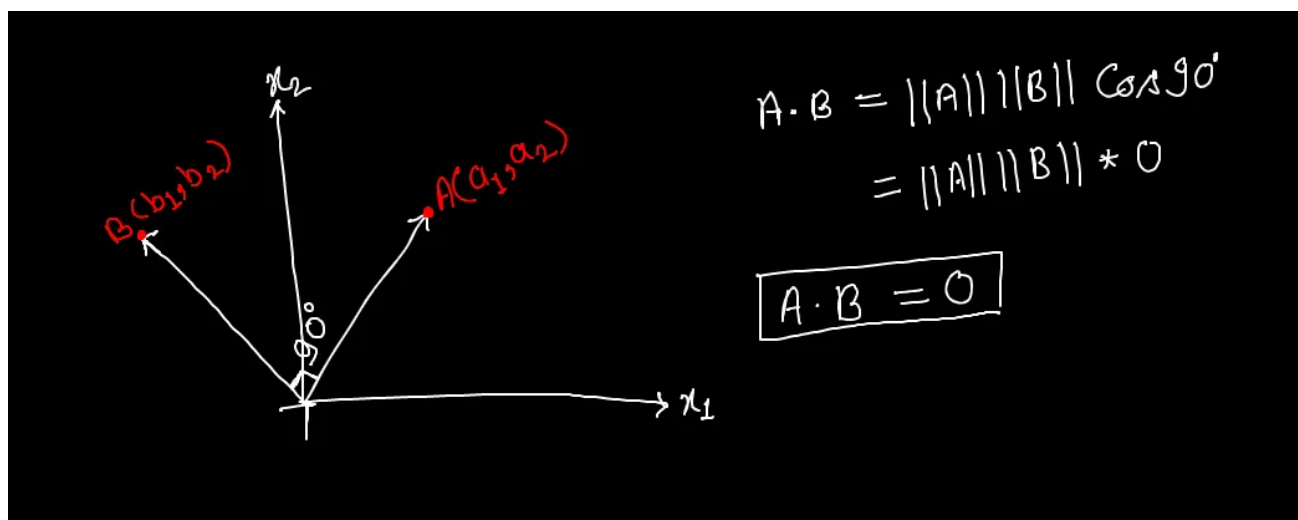$$A \cdot B = a_1 b_1 + a_2 b_2 = \|A\| \|B\| \cos \theta$$

$$\theta = \cos^{-1} \left\{ \frac{a_1 b_1 + a_2 b_2}{\|A\| \|B\|} \right\}$$

$$\|A\| = \sqrt{a_1^2 + a_2^2}$$

Angle between two vectors A and B

In the above image we can see how easily we can compute the angle between two vectors. Now let's look at one interesting case.

What if the dot product between two vectors is zero?

$$A \cdot B = \|A\| \|B\| \cos 90°$$
$$= \|A\| \|B\| * 0$$
$$\boxed{A \cdot B = 0}$$

Vectors Perpendicular to Each Other

Projection of one vector on another vector is like throwing light on one vector and projecting it's shadow on another vector. Let's see how to get projection of one vector on another vector.

In the above image we can see that if the dot product of two vectors is zero, both the vectors are perpendicular to each other.
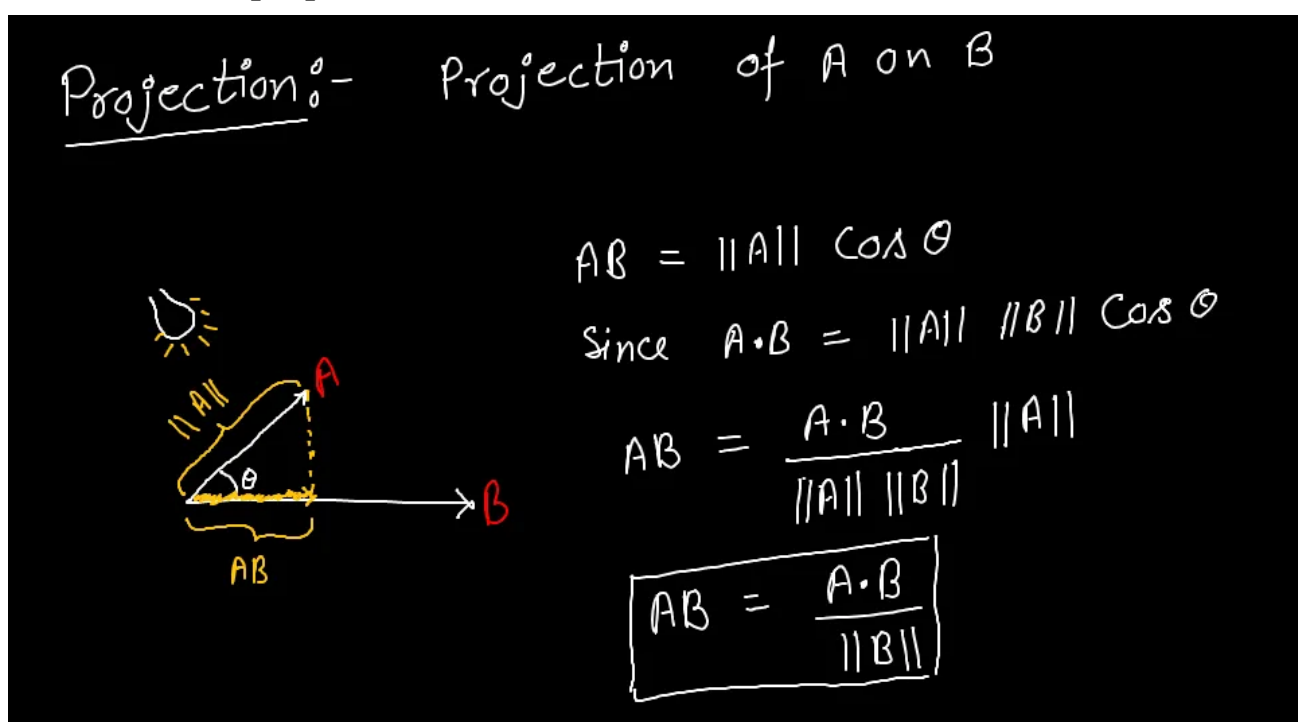

Projection of Vector A on B

In the above image AB is the projection of vector A on vector B. Just assume that a light is coming from the bulb above vector A and the shadow of A is getting projected on B.

Unit Vector:

A unit vector is represented by hat on top of vector. It represents the single unit of a vector.

A unit vector always has the same direction as the vector.

Length of the unit vector is 1, $\|A^\wedge\| = 1$
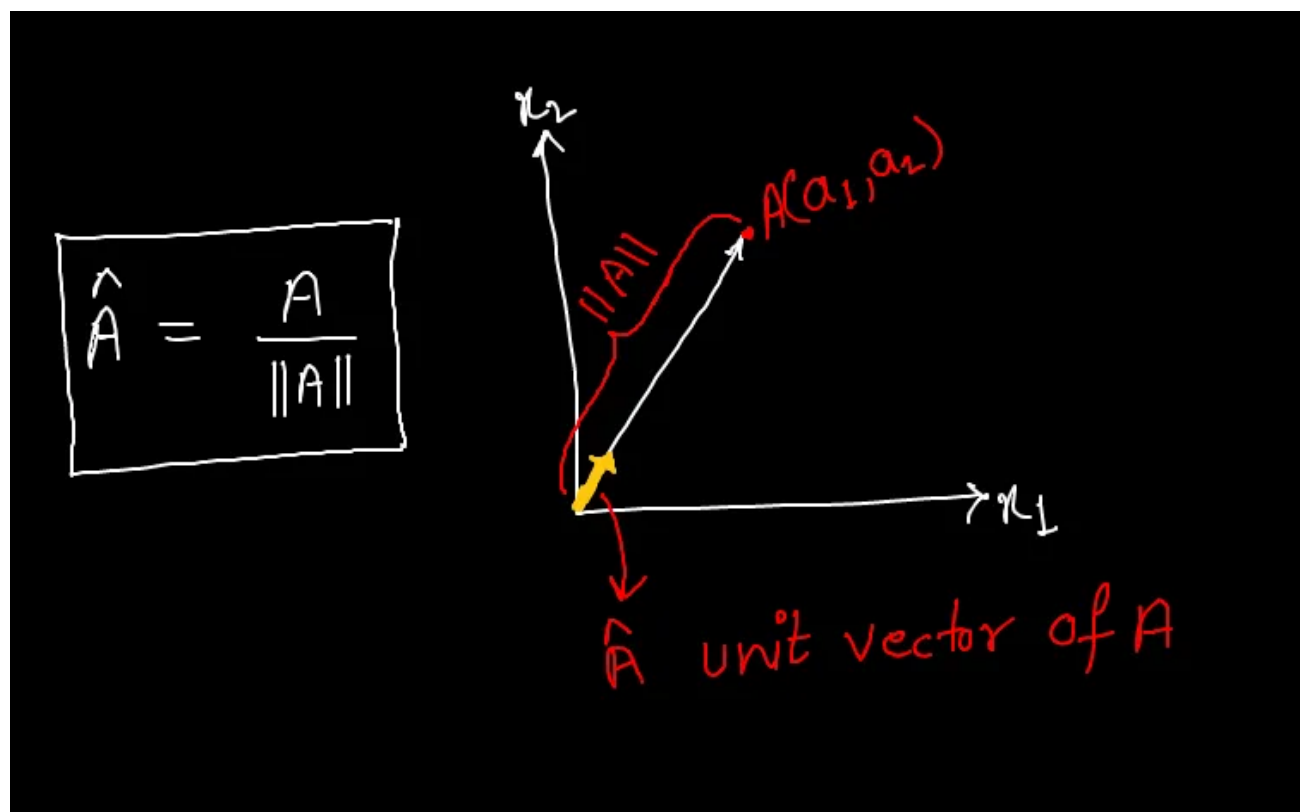
Unit Vector of Vector A

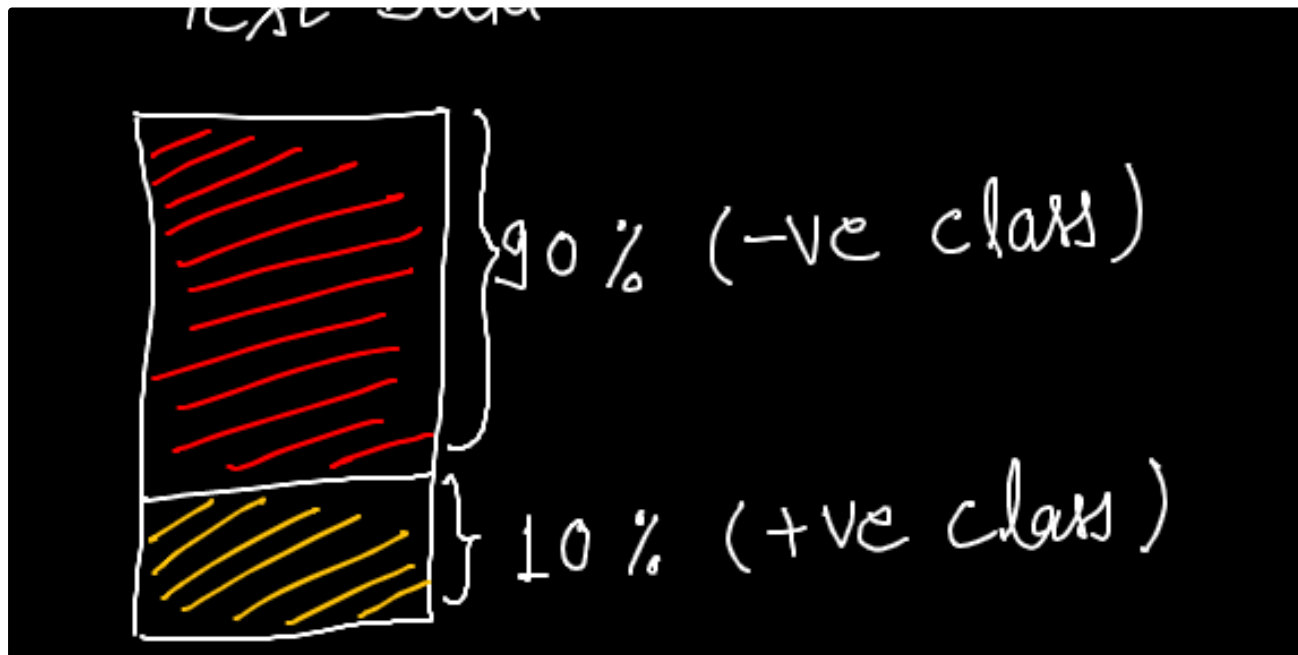Data Science   Linear Algebra   Machine Learning   Mathematics   Vector

Follow

Written by Vishal Kumar

30 Followers

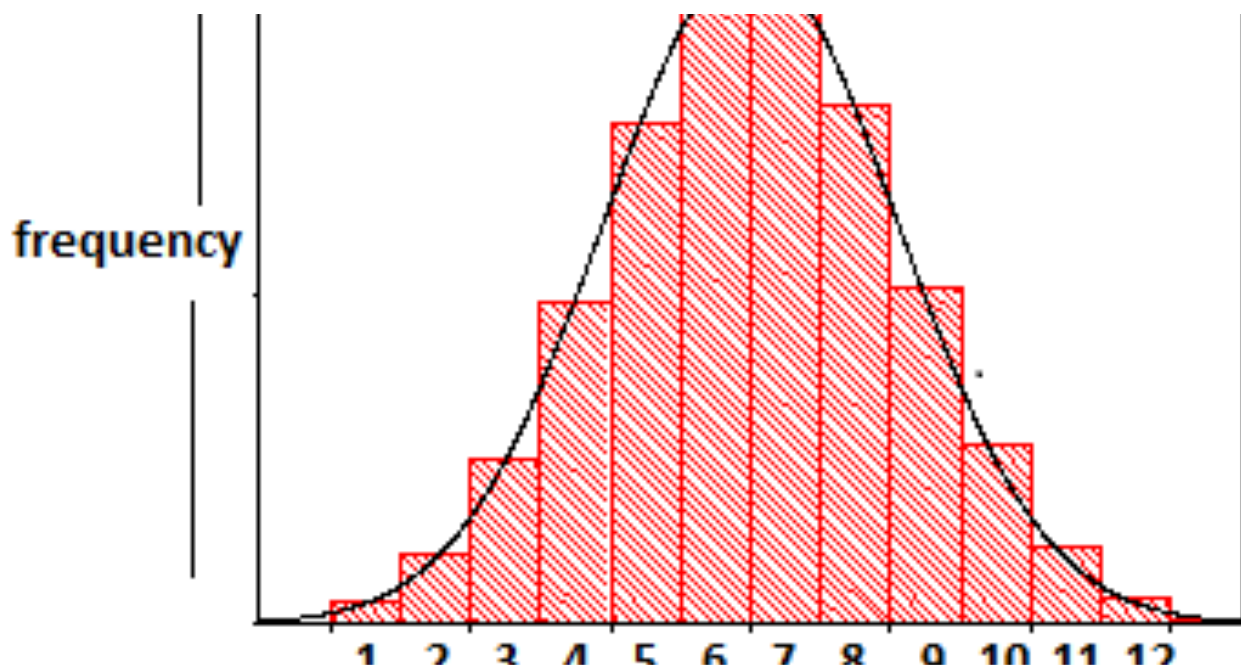Data Scientist, Data Science Enthusiast working on NLP, Knowledge Graphs and deep learning.

More from Vishal Kumar

Vishal Kumar

Precision or Recall When to Choose What? Can F1-Score help?

Is Accuracy a good measure?

Jul 7, 2022   👋 6

Vishal Kumar

Measurement of Dispersion of Data using Range, Variance and Standard Deviation

In data science domain when we see data we first try to measure the central tendency of the data then in second step we try to measure the...
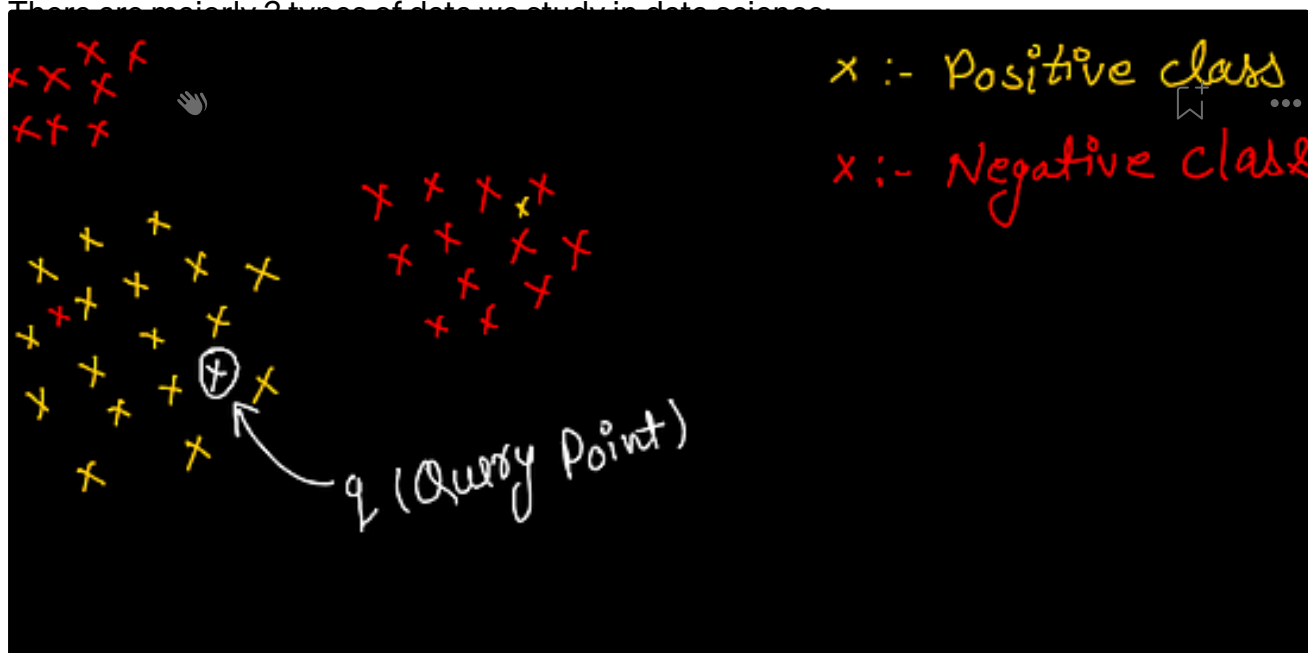
Jul 17, 2021   👏 110   💬 1

![Vishal Kumar] Vishal Kumar

Types of DATA! in Data Science

There are majorly 3 types of data we study in data science:

![Vishal Kumar] Vishal Kumar

K-Nearest Neighbors A Simple Yet Powerful Machine Learning Algorithm

Check the video explanation of this article here.

Apr 22, 2022   👋 6

Recommended from Medium

Shaw Talebi in Towards Data Science

5 AI Projects You Can Build This Weekend (with Python)
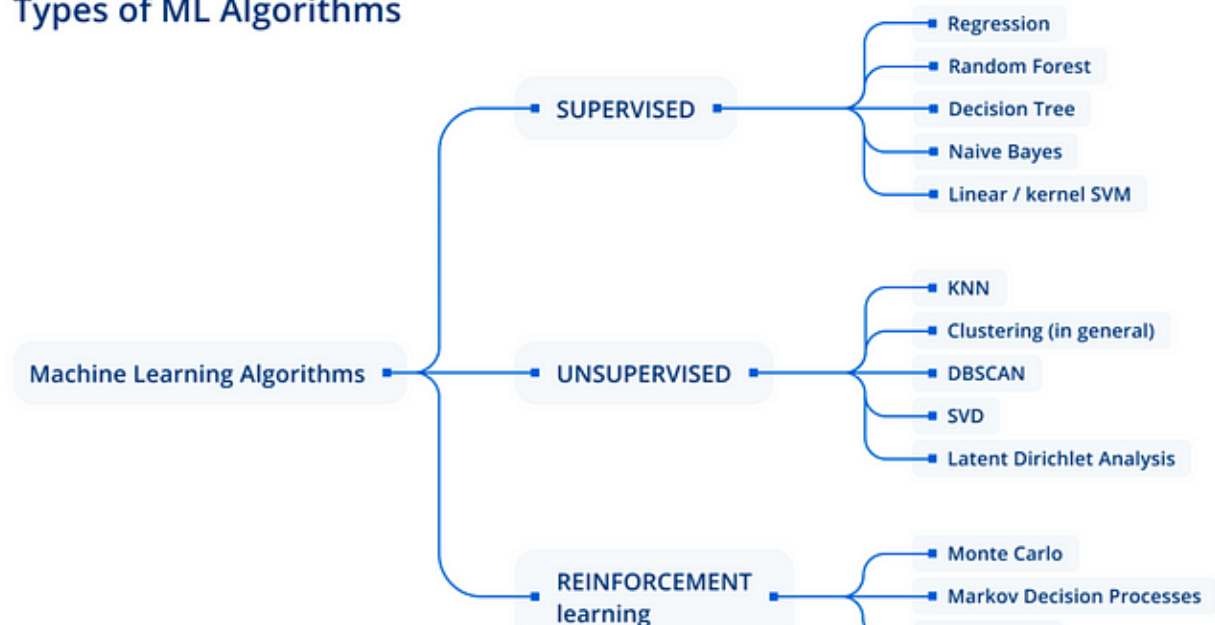
From beginner-friendly to advanced

✦   Oct 9   👏 2.4K   💬 32   ⊓⁺   •••

Types of ML Algorithms

John Vastola

10 Must-Know Machine Learning Algorithms for Data Scientists

Machine learning is the science of getting computers to act without being explicitly programmed." — Andrew Ng
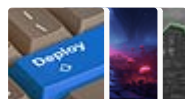
✦   Dec 6, 2022   👏 473   💬 7   ⊓⁺   •••

Lists

Predictive Modeling w/ Python
20 stories · 1596 saves

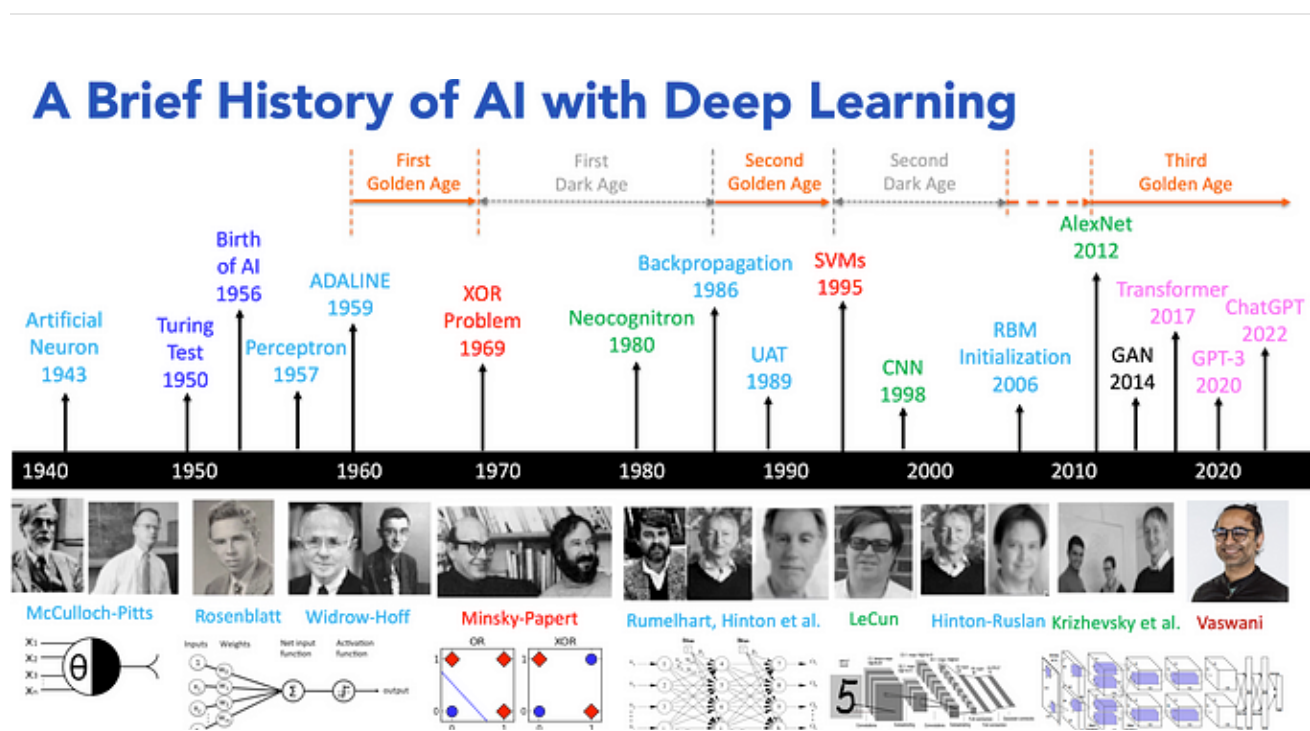**Practical Guides to Machine Learning**
10 stories · 1941 saves

Natural Language Processing
1758 stories · 1355 saves
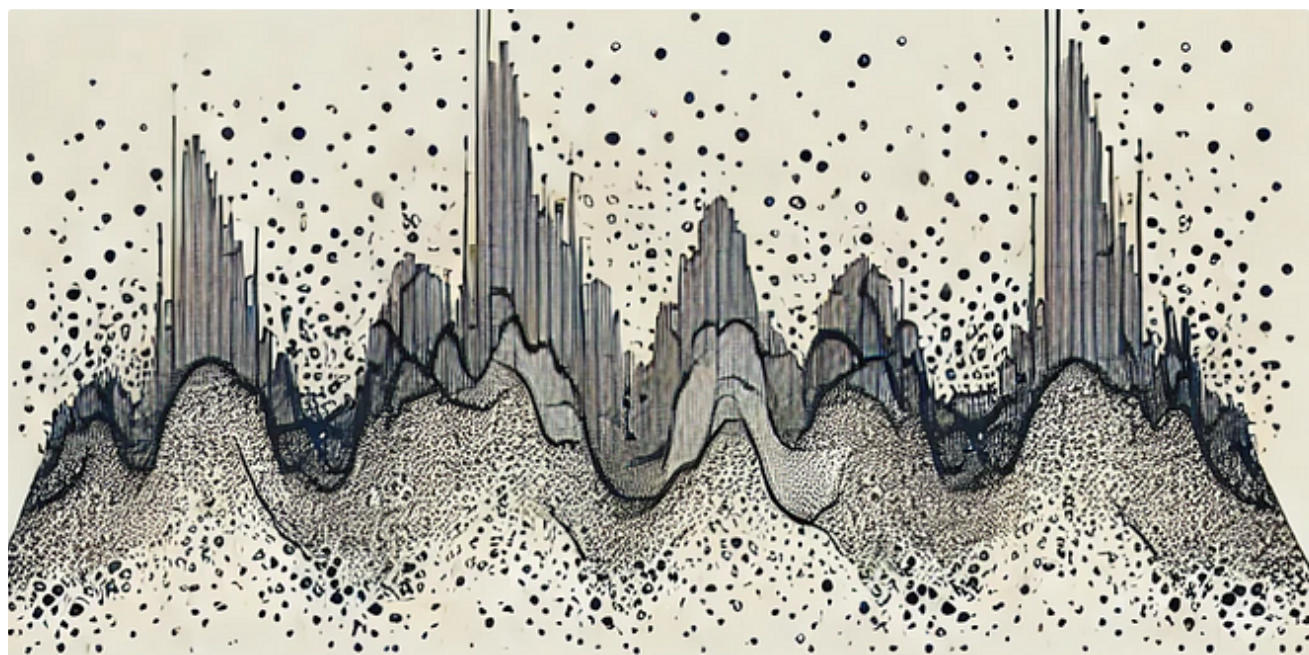
data science and AI
40 stories · 265 saves

A Brief History of AI with Deep Learning

Artificial intelligence (AI) and deep learning have seen remarkable progress over the past several decades, transforming fields like...
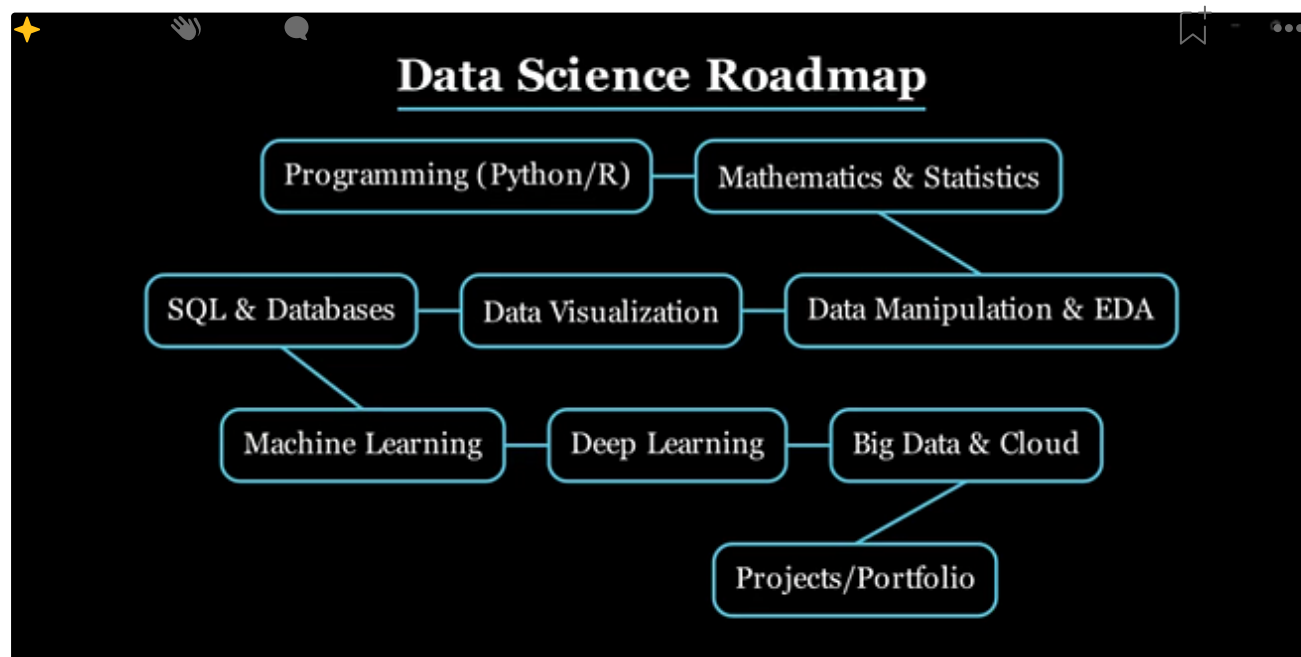
✦ Sep 1   👋 187   💬 3   🔖+   •••

H Heiko Limberg

Smoothen out Nearest Neighbor Regression

Sometimes it is hard to find THE function which fits your data. Maybe because the data behaves differently in different regions or we...
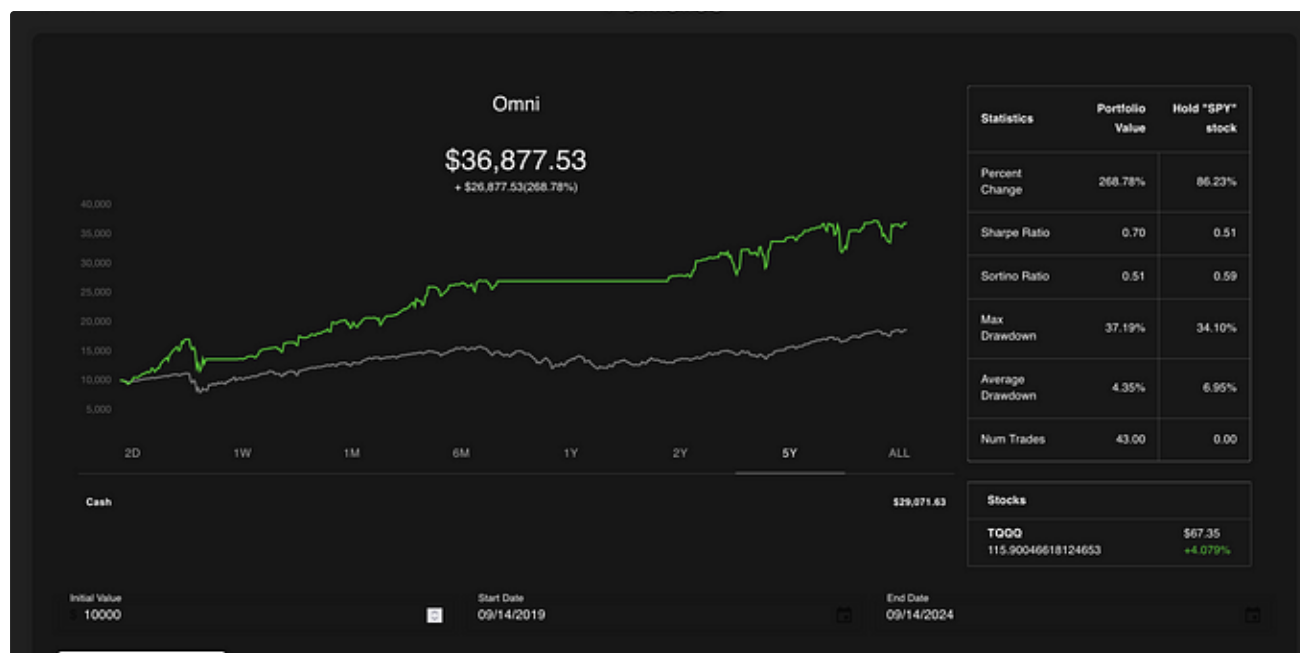

📖 Daily Deep learning

🎯 So You Want to Be a Data Scientist? Here's Your Roadmap to Success!

So, you've heard the buzz — Data Science is one of the hottest career paths out there. You're intrigued, maybe even excited about the idea...

Sep 24

Austin Starks in DataDrivenInvestor

I used OpenAI's o1 model to develop a trading strategy. It is DESTROYING the market

It literally took one try. I was shocked.

★ Sep 15 👋 3.7K 💬 104   🔖⁺   •••

See more recommendations