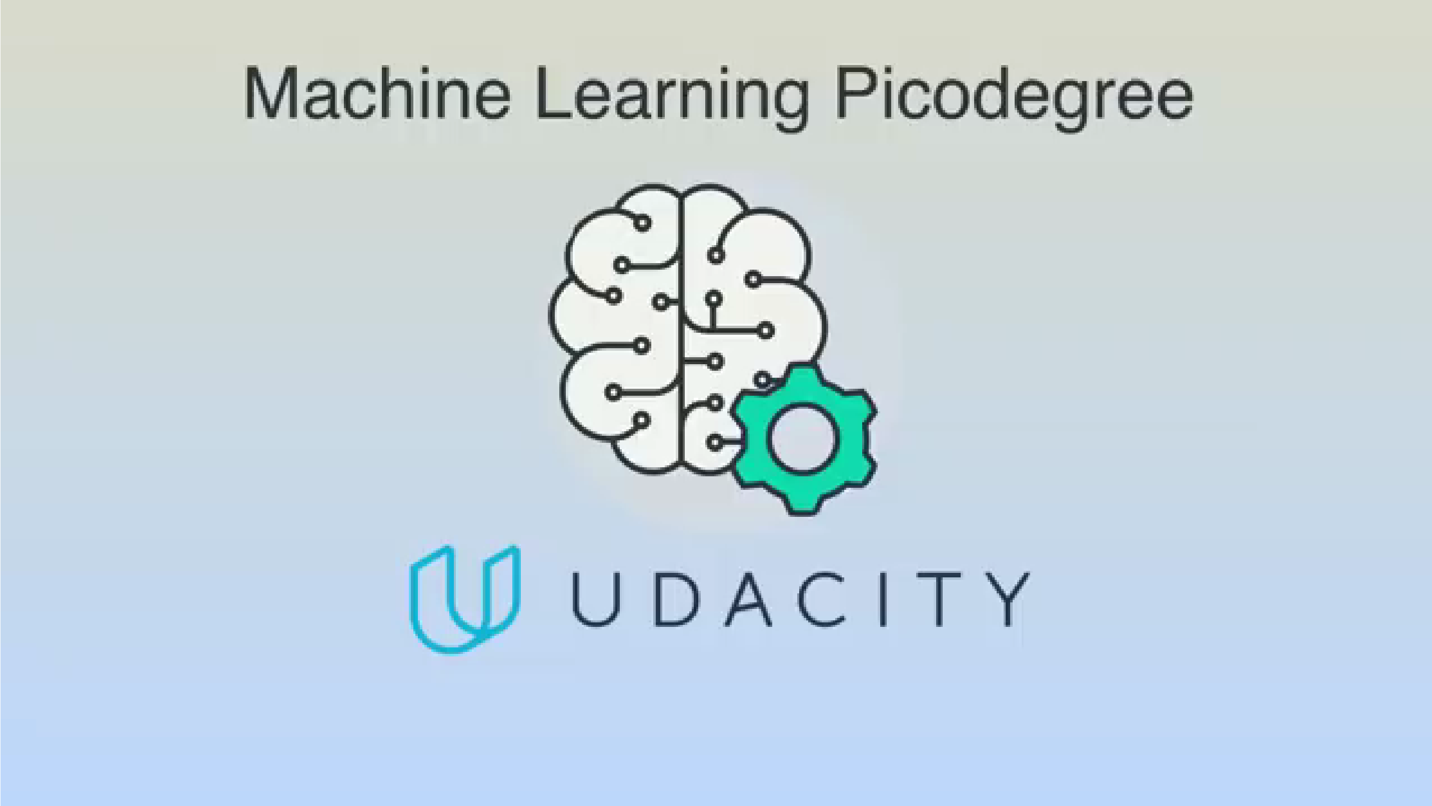
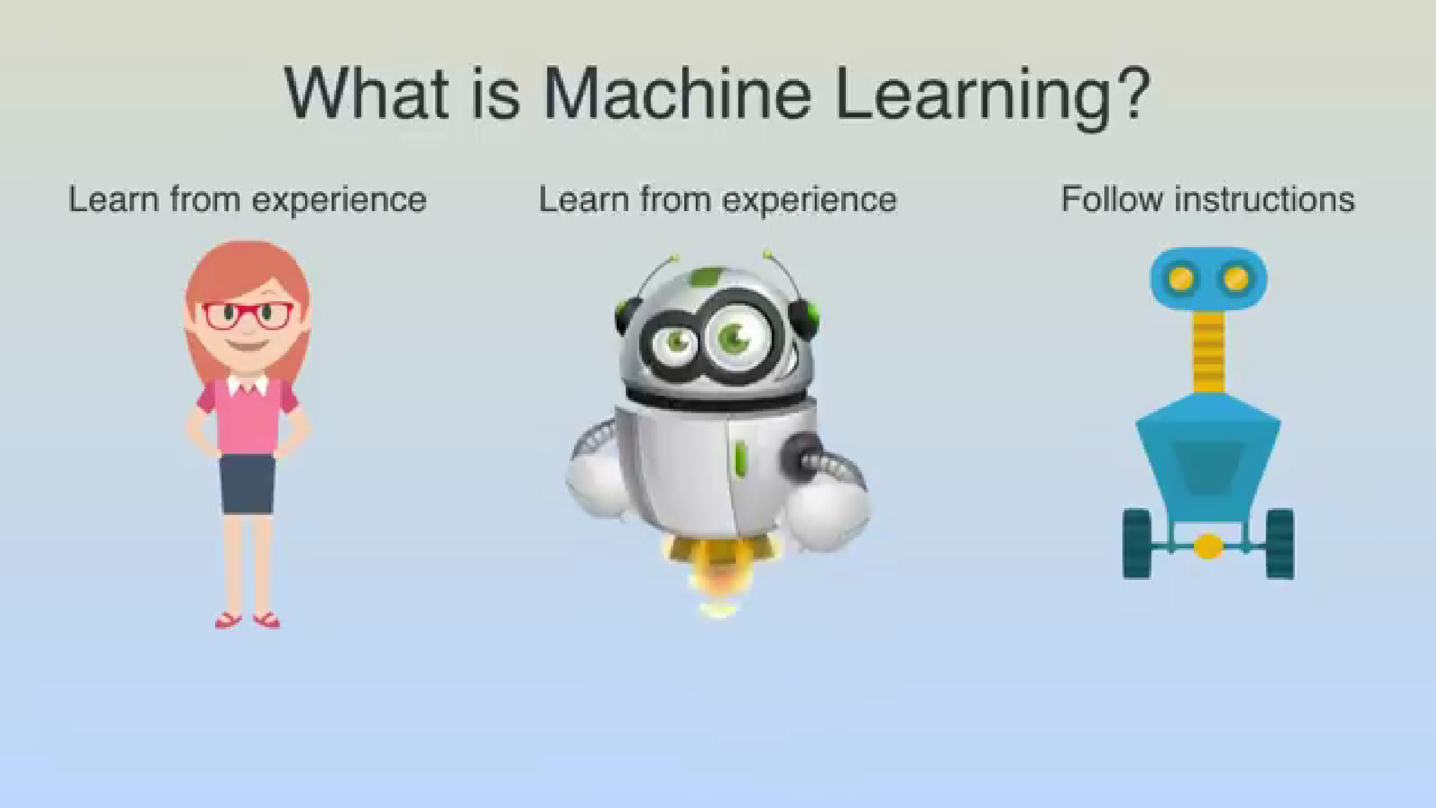
(137) A Friendly Introduction to Machine Learning - YouTube



Machine learning is a field that aims to bridge the gap between humans and computers. Unlike humans who learn from past experiences, computers traditionally follow programmed instructions. However, through machine learning, computers can also learn from experience. This is the focus of Udasi's P degree in machine learning.

(137) A Friendly Introduction to Machine Learning - YouTube



Machine learning is essentially teaching computers to learn from past data. This process can be simple and is not something to be feared. For instance, it can be used to analyze the housing market and predict housing prices.

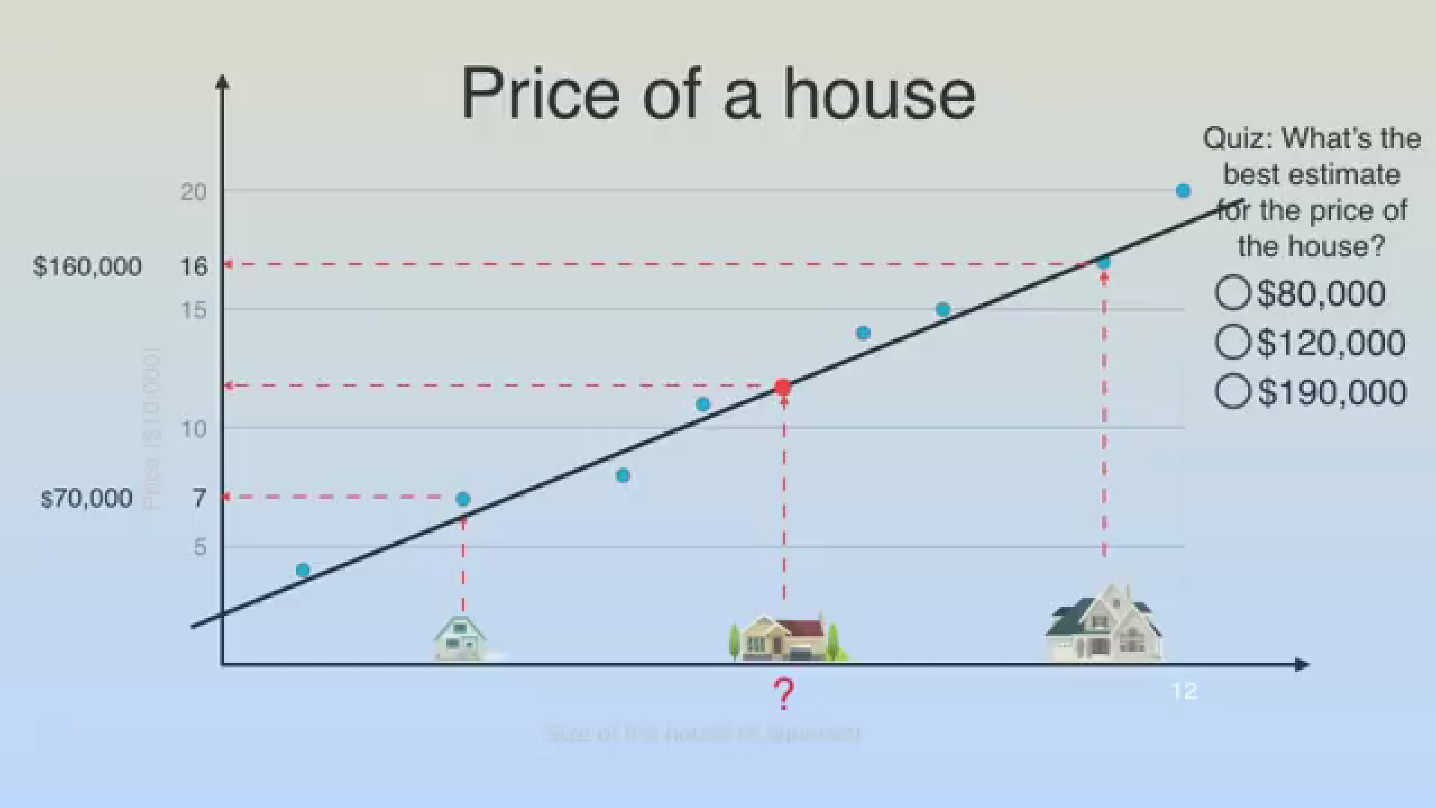
(137) A Friendly Introduction to Machine Learning - YouTube

The lecture discusses estimating the price of a medium-sized house using a grid system. The x-axis represents the house size in square feet, while the y-axis represents the price. This method utilizes data from previously sold houses, represented by blue dots.

(137) A Friendly Introduction to Machine Learning - YouTube

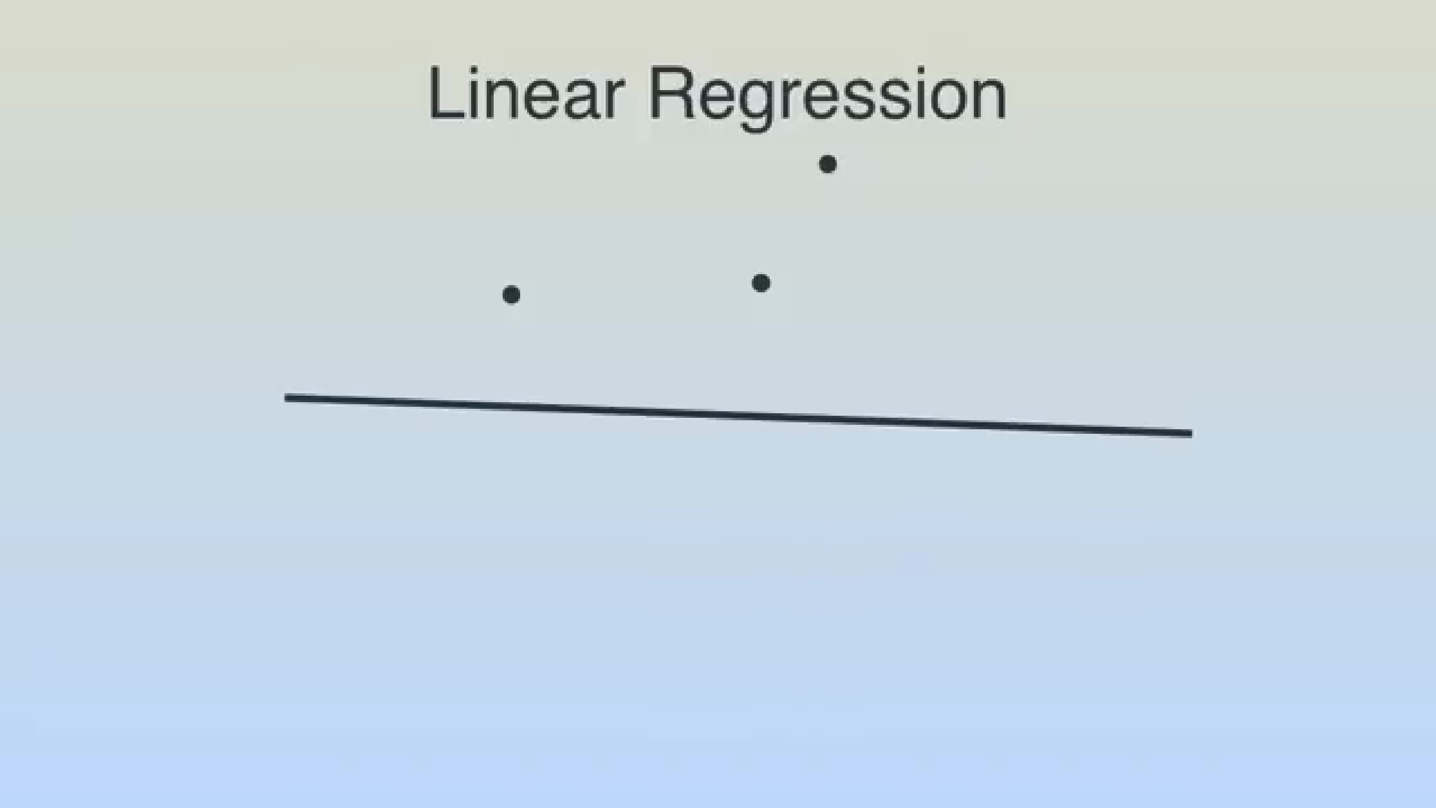
The graph indicates a small house costs $70,000 and a big house is $160,000. A quiz asks for the best guess of a medium house's price. By drawing a line that fits the data, the best guess for the medium house's price can be determined.

(137) A Friendly Introduction to Machine Learning - YouTube



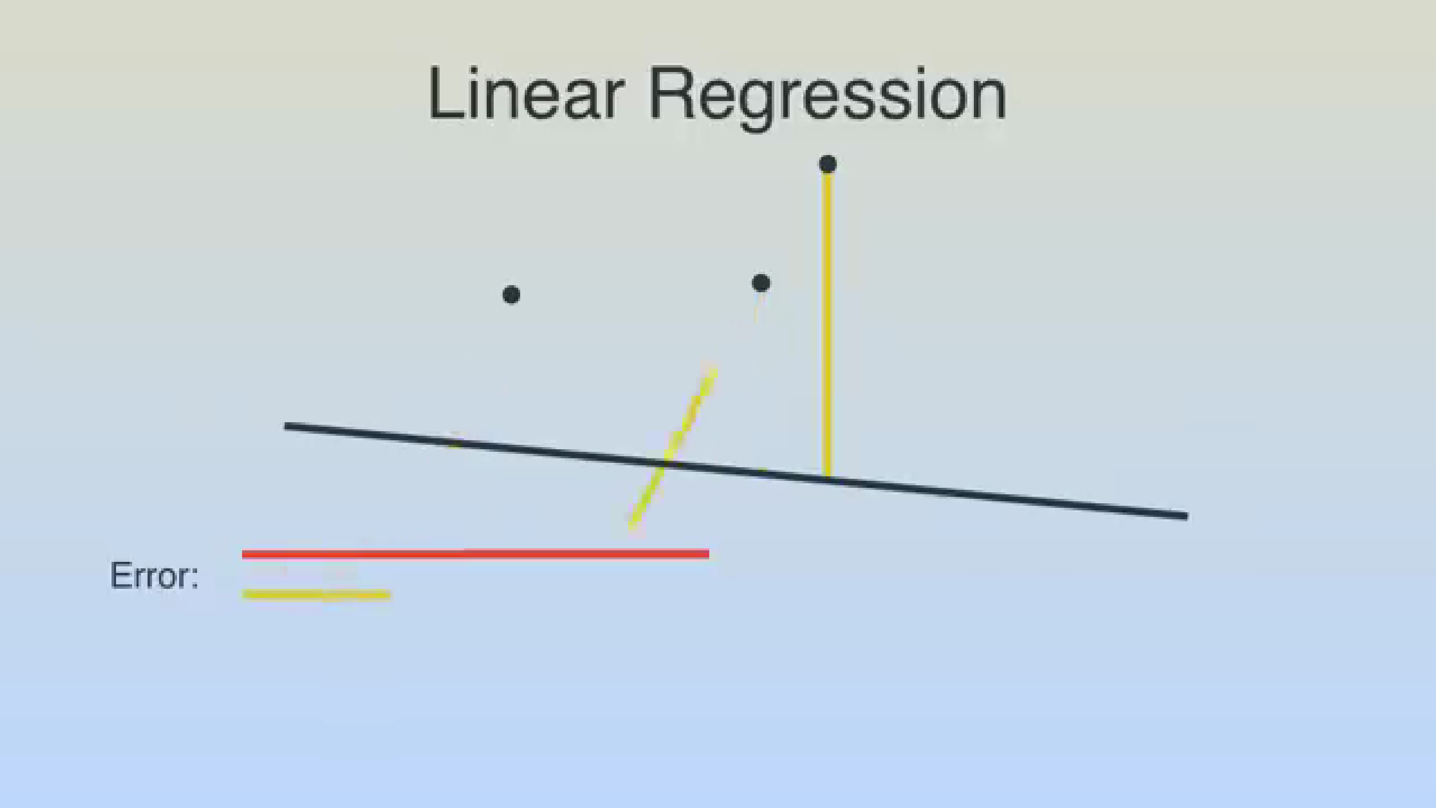
Linear regression is a method used to find the best fitting line through data points. In this context, the best line is subjective. The process involves teaching the computer to draw a random line through the points and then evaluate its accuracy.

(137) A Friendly Introduction to Machine Learning - YouTube



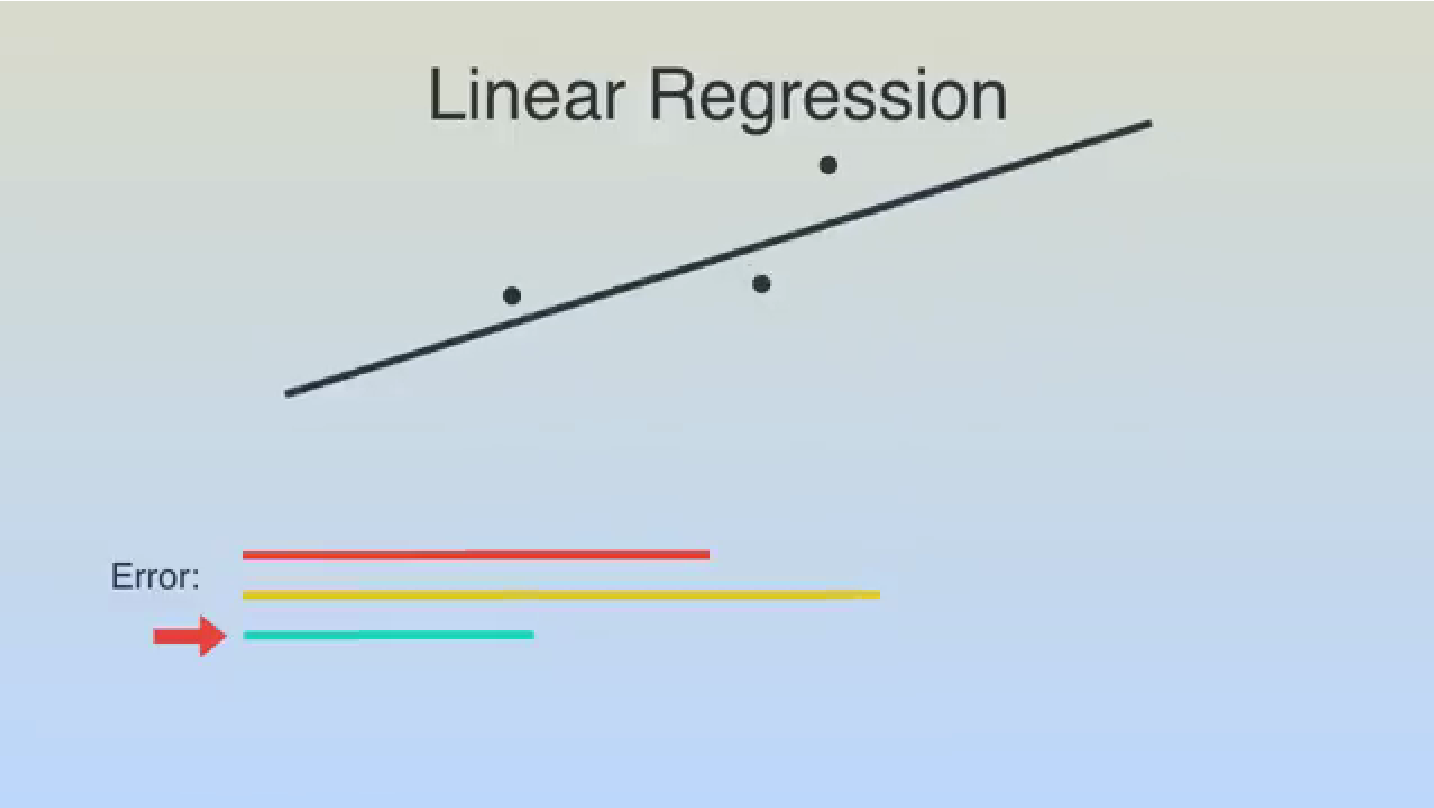
To assess the accuracy of a line, the error is calculated by measuring the distances from the line to three points. This error, represented by the sum of these distances, can be minimized by adjusting the line's position. The new error is then calculated using the same method.

(137) A Friendly Introduction to Machine Learning - YouTube



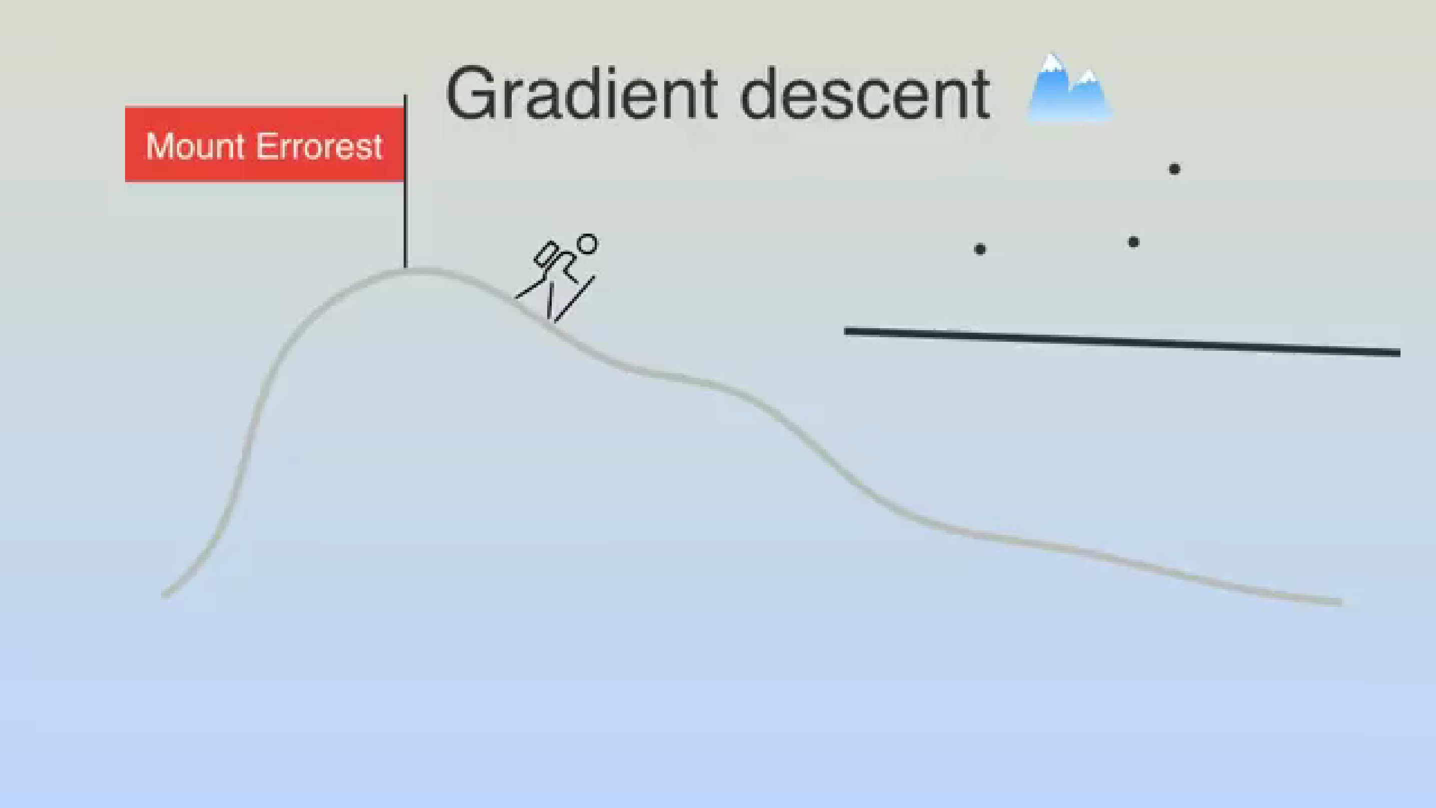
The process of gradient descent involves adjusting variables to minimize error. By moving in the opposite direction of initial error, recalculating, and repeating this process, the error gradually decreases. This iterative method eventually leads to an optimal solution, represented by a line.

(137) A Friendly Introduction to Machine Learning - YouTube



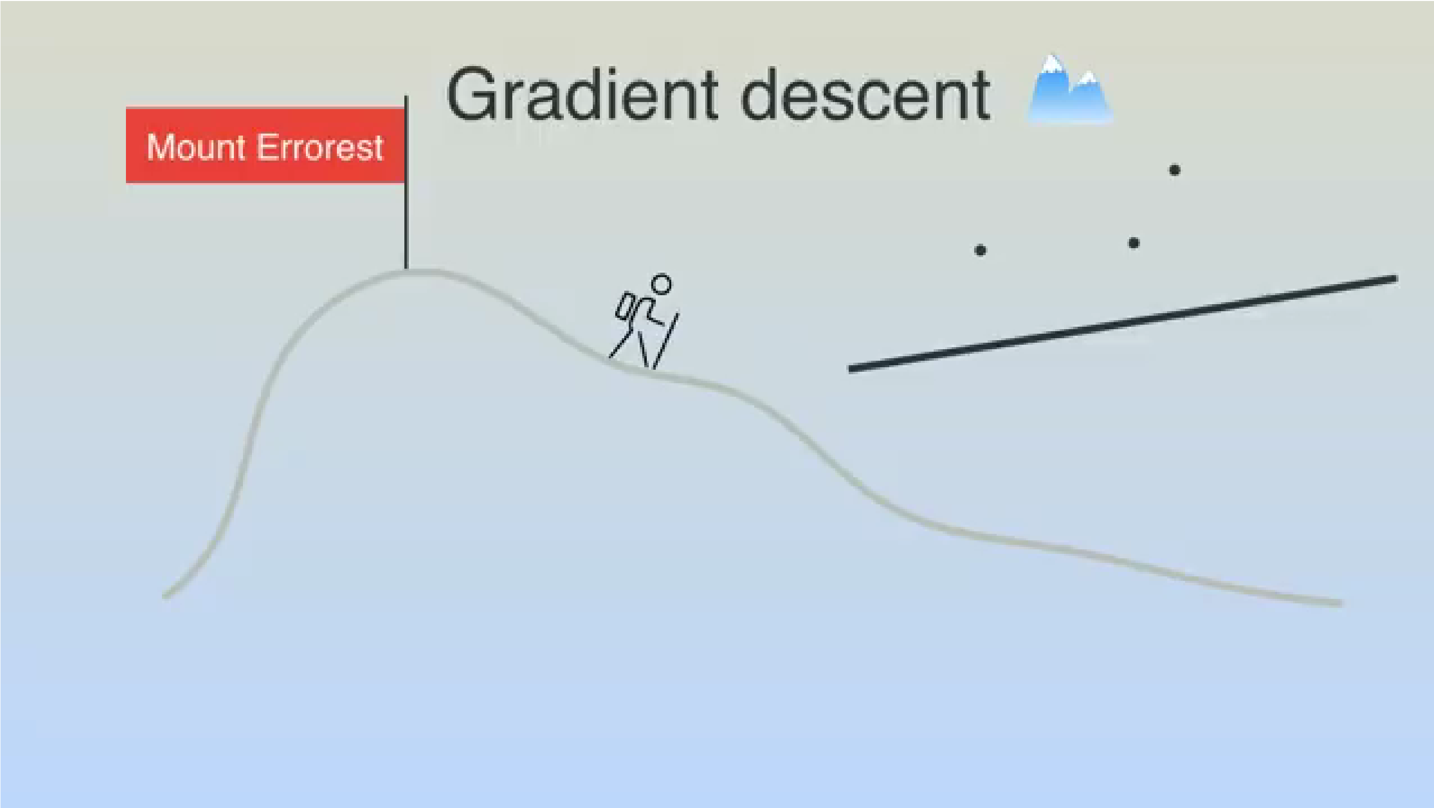
In real-life applications, negative distances are not preferred, so the square of the distance from the point to the line is used instead, a method known as least squares. This is likened to descending a mountain, where the goal is to reduce the error, similar to lowering altitude.

(137) A Friendly Introduction to Machine Learning - YouTube



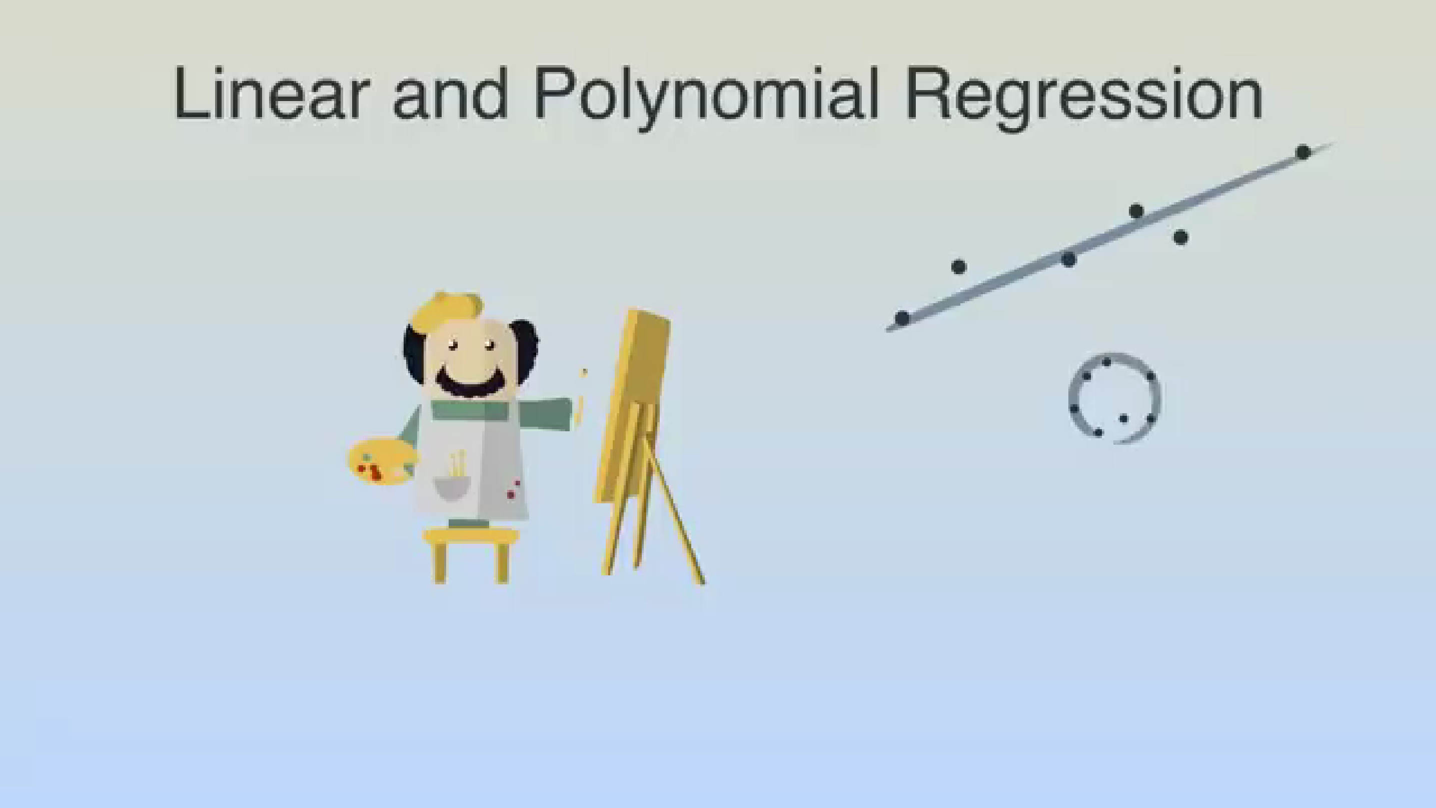
The lecture discusses the process of error minimization in a given environment. It explains the concept by comparing the leftward movement, which increases error, to descending and moving the line down. Conversely, moving right decreases error, akin to ascending and moving the line up.

(137) A Friendly Introduction to Machine Learning - YouTube



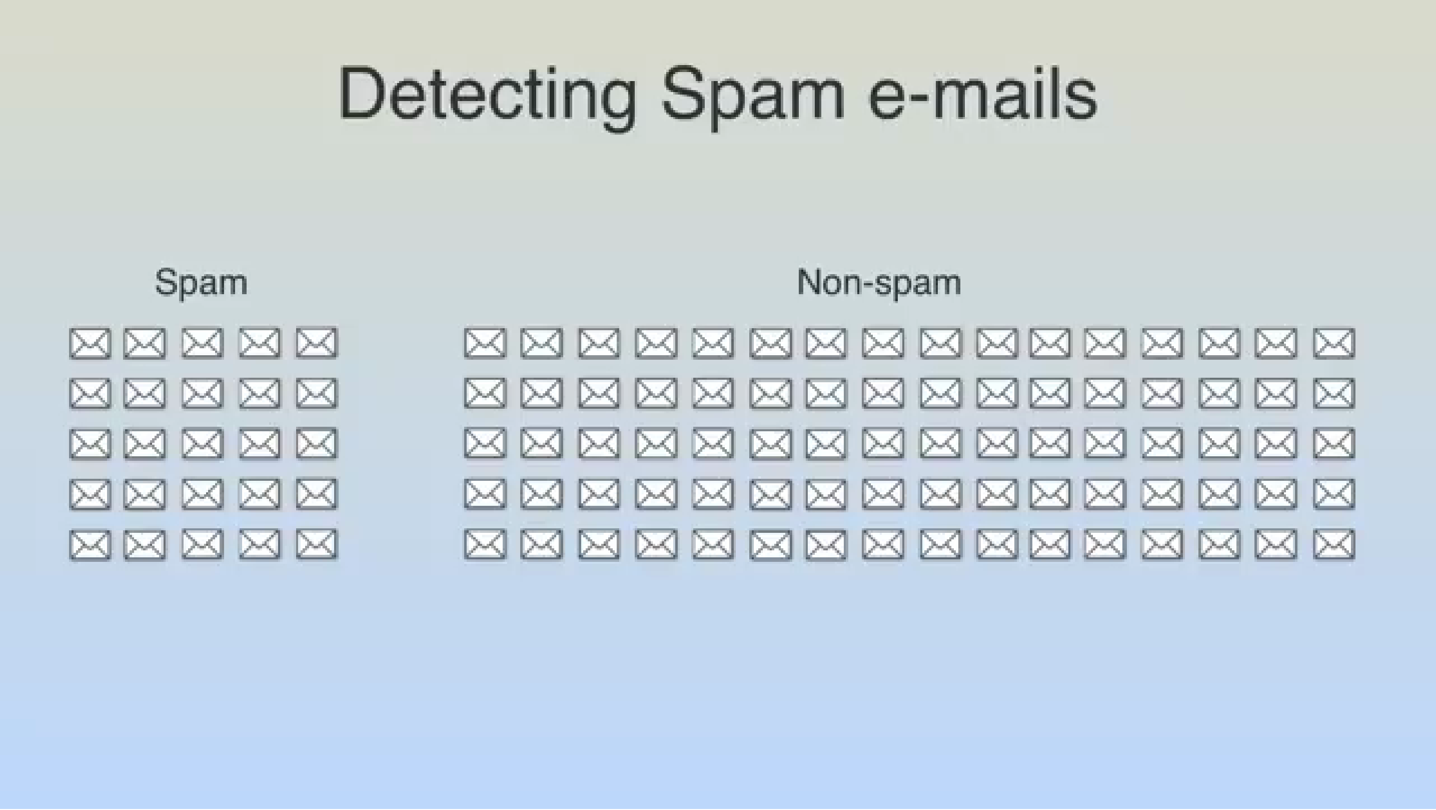
Linear regression can be compared to a painter drawing the best fitting line through data. The process is repeated until the error is minimized, symbolizing a successful descent from a mountain. This method can also fit a circle, parabola, or higher degree curve if the data doesn't form a straight line

(137) A Friendly Introduction to Machine Learning - YouTube



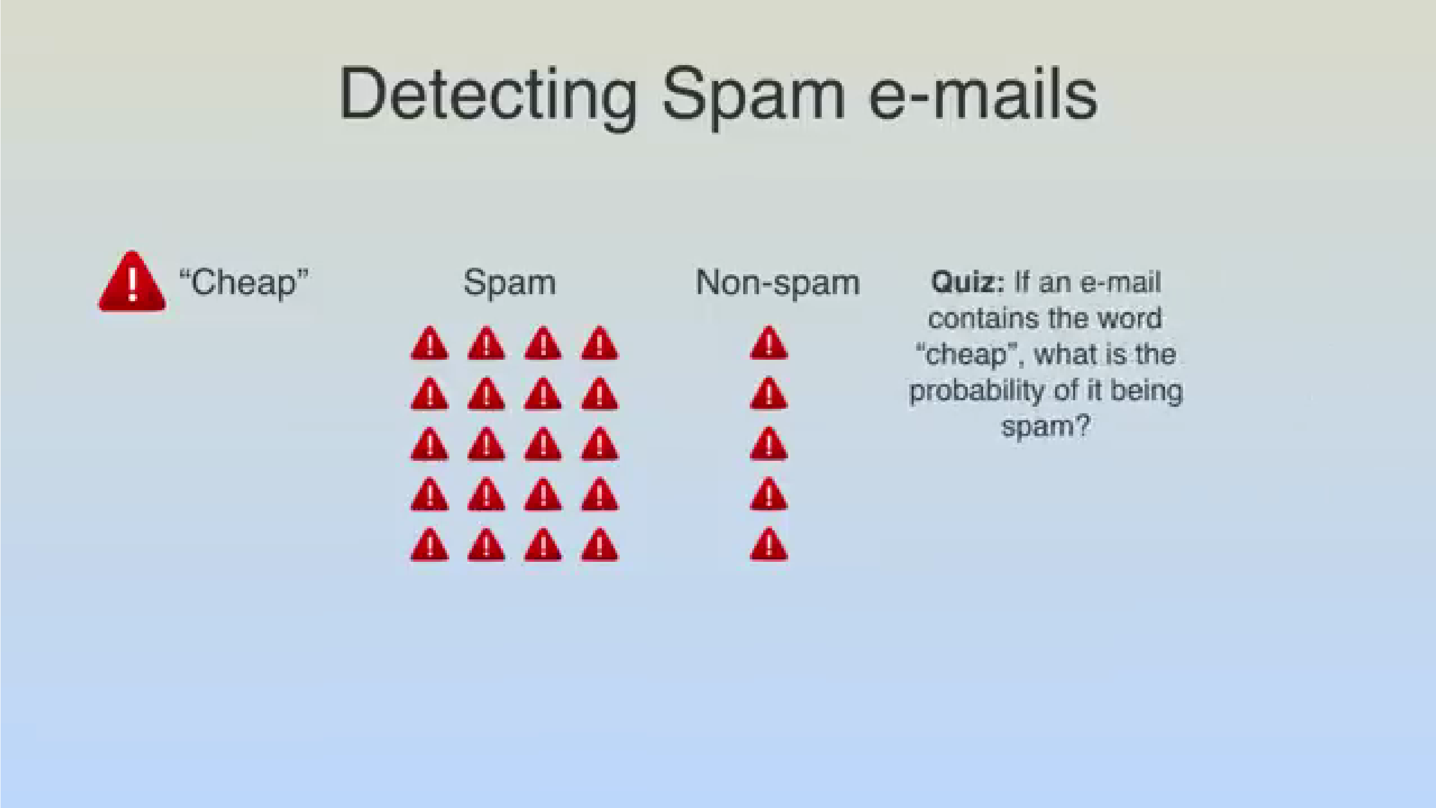
The text discusses fitting a cubic polynomial to data. It then transitions to building an email spam detection classifier, which is trained using a dataset of 100 emails, with 25 flagged as spam and 75 as non-spam. The classifier works by identifying features common in spam emails.

(137) A Friendly Introduction to Machine Learning - YouTube



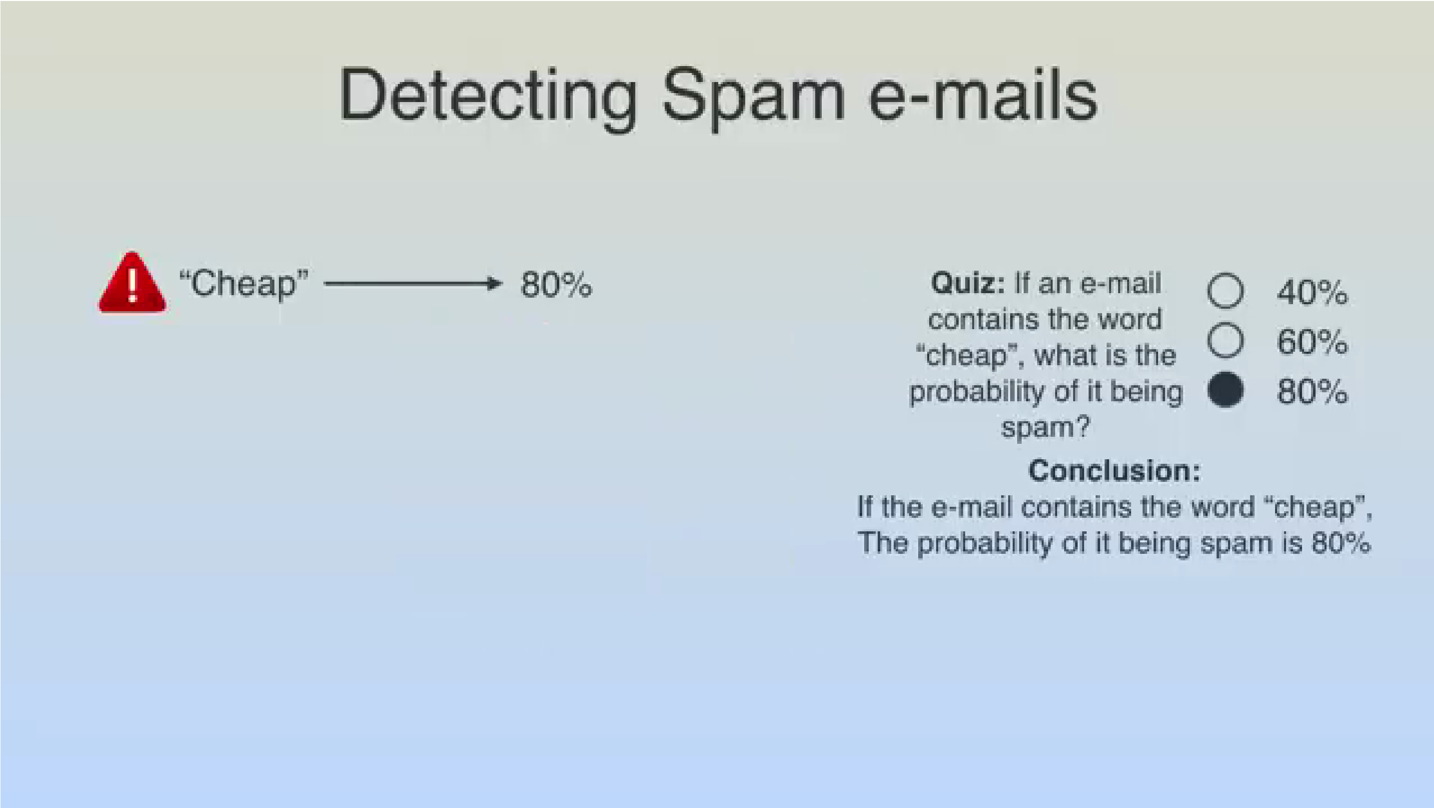
The text discusses how the word 'cheap' in an email could indicate spam. After analyzing 100 emails, it was found that 20 spam emails and five non-spam emails contained the word 'cheap'. This suggests a possible correlation between the word 'cheap' and spam emails.

(137) A Friendly Introduction to Machine Learning - YouTube



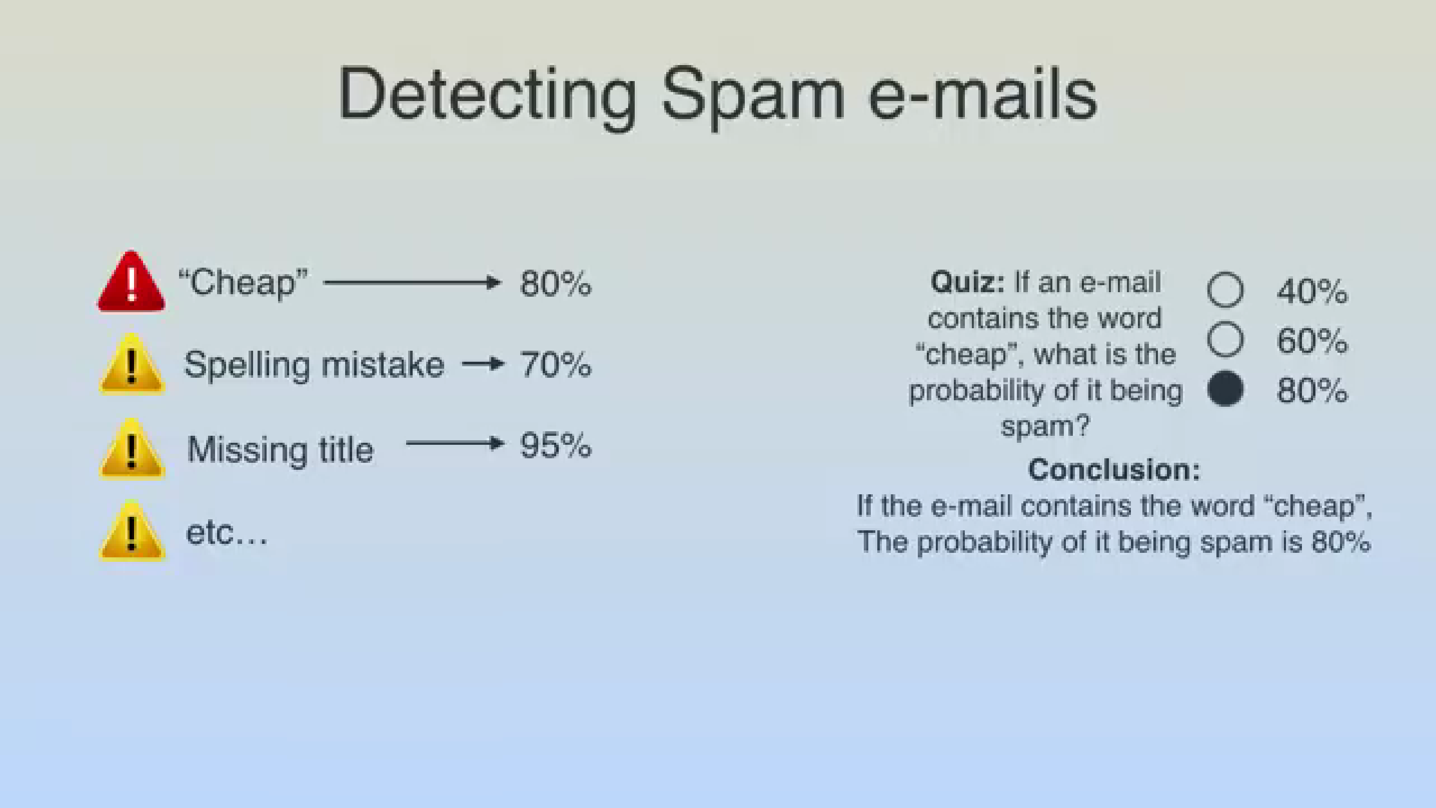
Analyzing data from 25 emails containing the word "cheap", it was found that 20 were spam, forming an 80/20 split. Thus, the probability of an email being spam if it contains the word "cheap" is 80%. This feature is now associated with an 80

(137) A Friendly Introduction to Machine Learning - YouTube



The algorithm flags future messages as spam by assessing certain features and their associated probabilities. For instance, emails with spelling mistakes have a 70% chance of being spam, while those missing a title have a 95% chance. These features are combined to predict if future emails are spam.

(137) A Friendly Introduction to Machine Learning - YouTube

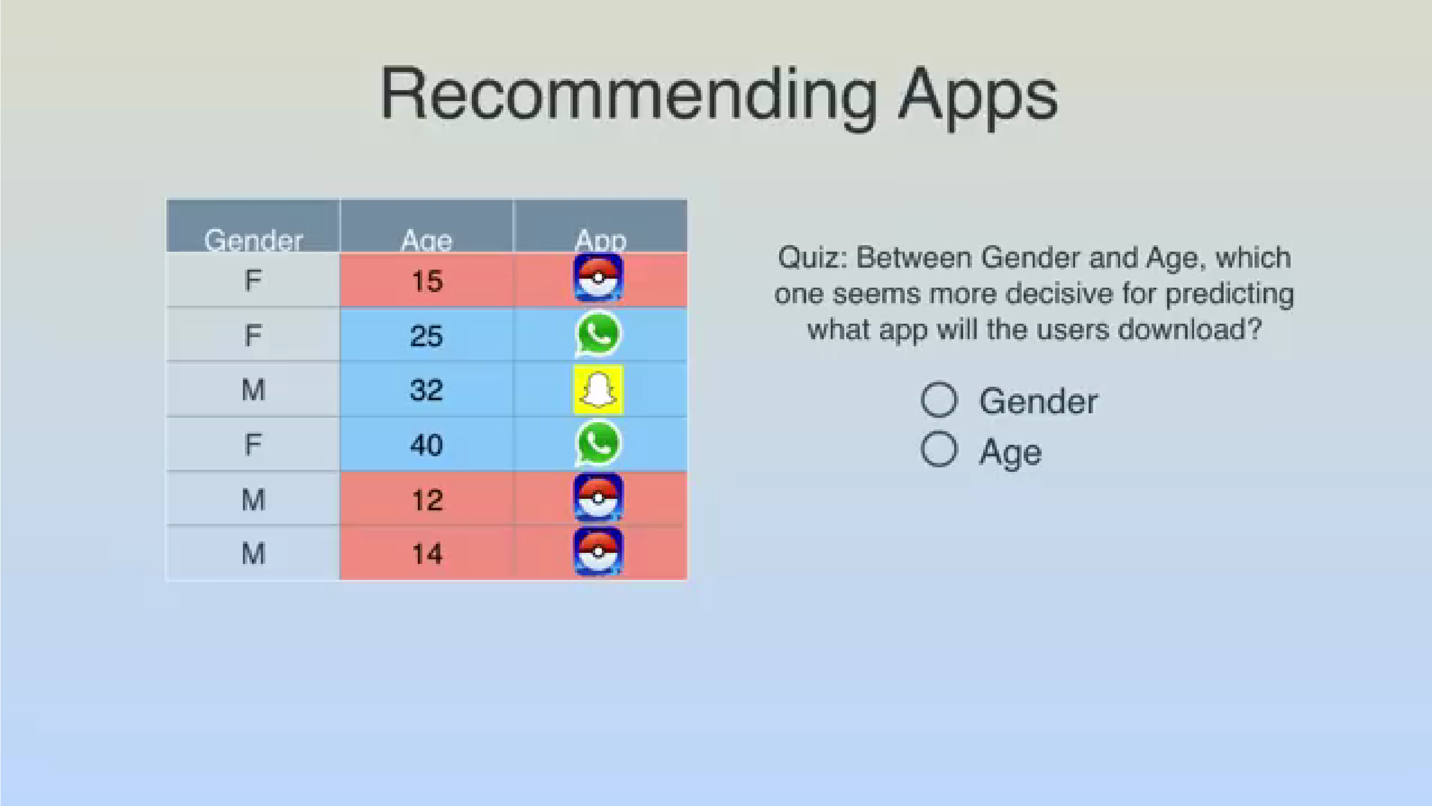


The Naive Based Algorithm is used to recommend apps to users on platforms like App Store or Google Play. Recommendations are made based on a data table containing user information such as gender, age, and previously downloaded apps, aiming to predict the apps users are most likely to download.

(137) A Friendly Introduction to Machine Learning - YouTube

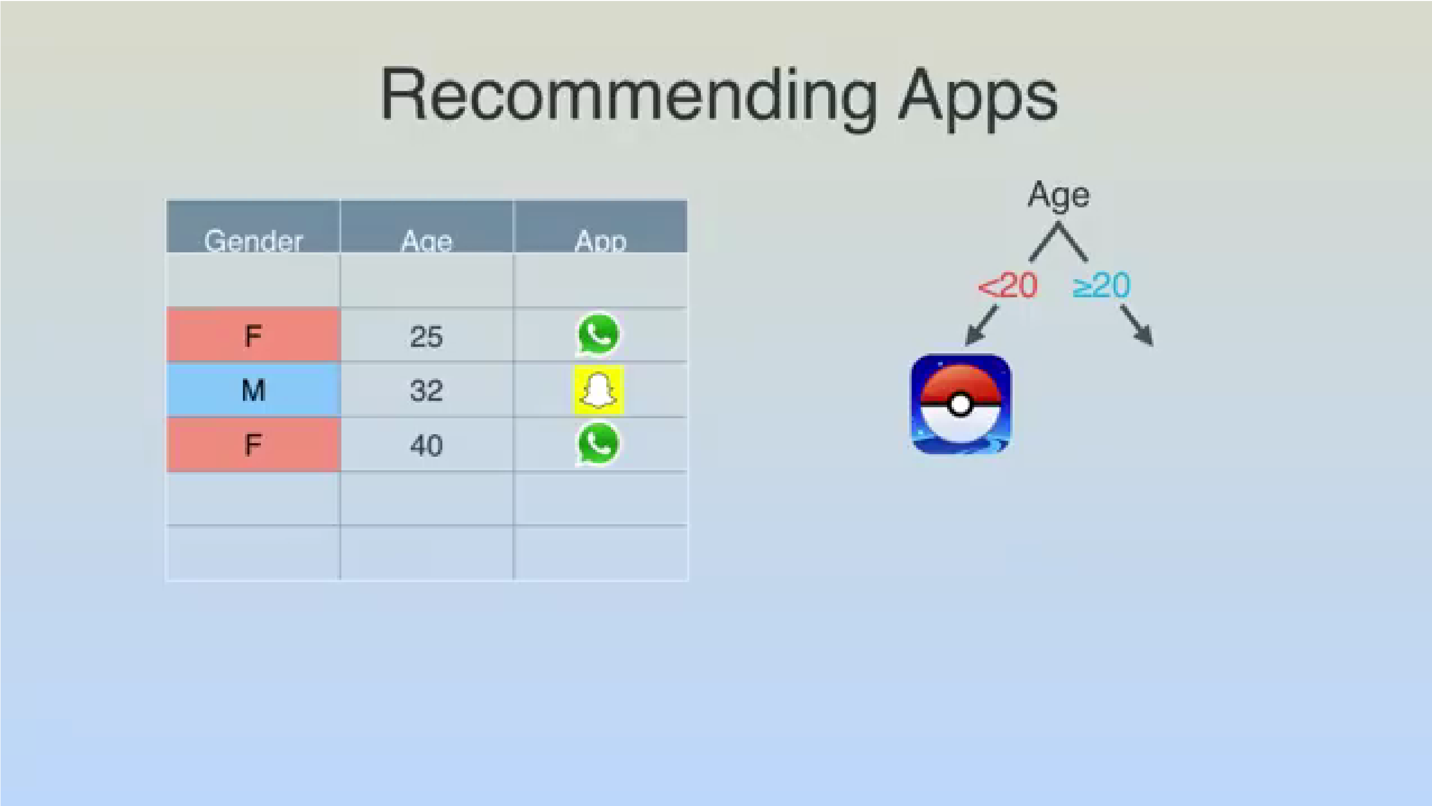
The choice of app downloaded appears more dependent on age than gender. Regardless of gender, Pokémon Go was downloaded by all users under 20 years old, while those aged 20 and above did not. Gender showed less influence, with both males and females downloading Pokémon Go, albeit on different platforms.

(137) A Friendly Introduction to Machine Learning - YouTube



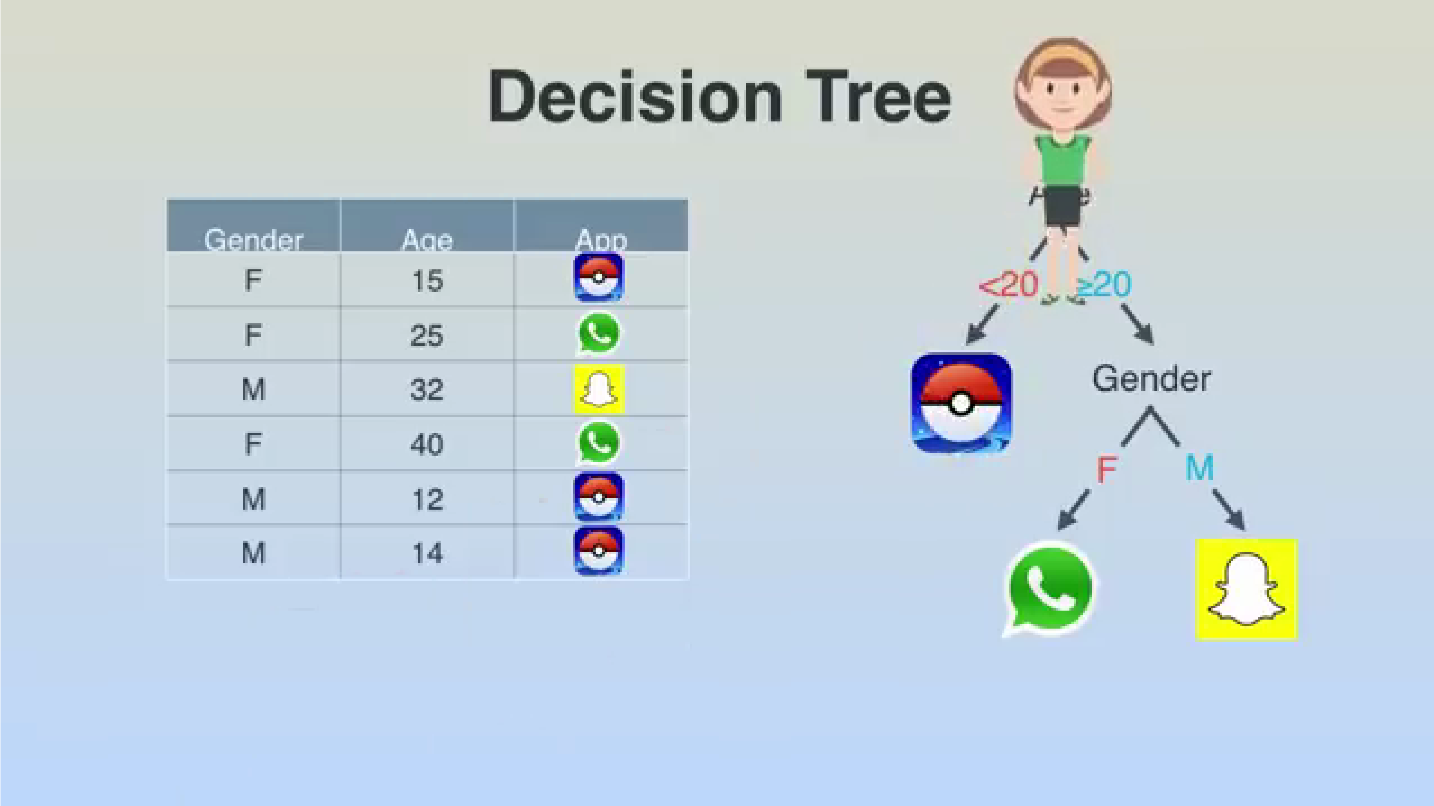
The lecture discussed using age as a primary data splitter. If a person is under 20, Pokémon Go is recommended. If they're 20 or older, recommendations are based on gender: WhatsApp for females and Snapchat for males.

(137) A Friendly Introduction to Machine Learning - YouTube



The text discusses a decision tree model that recommends apps based on user demographics. For instance, WhatsApp is suggested for females, Snapchat for males, Pokémon Go for younger users, and other apps are recommended for older users after additional checks.

(137) A Friendly Introduction to Machine Learning - YouTube

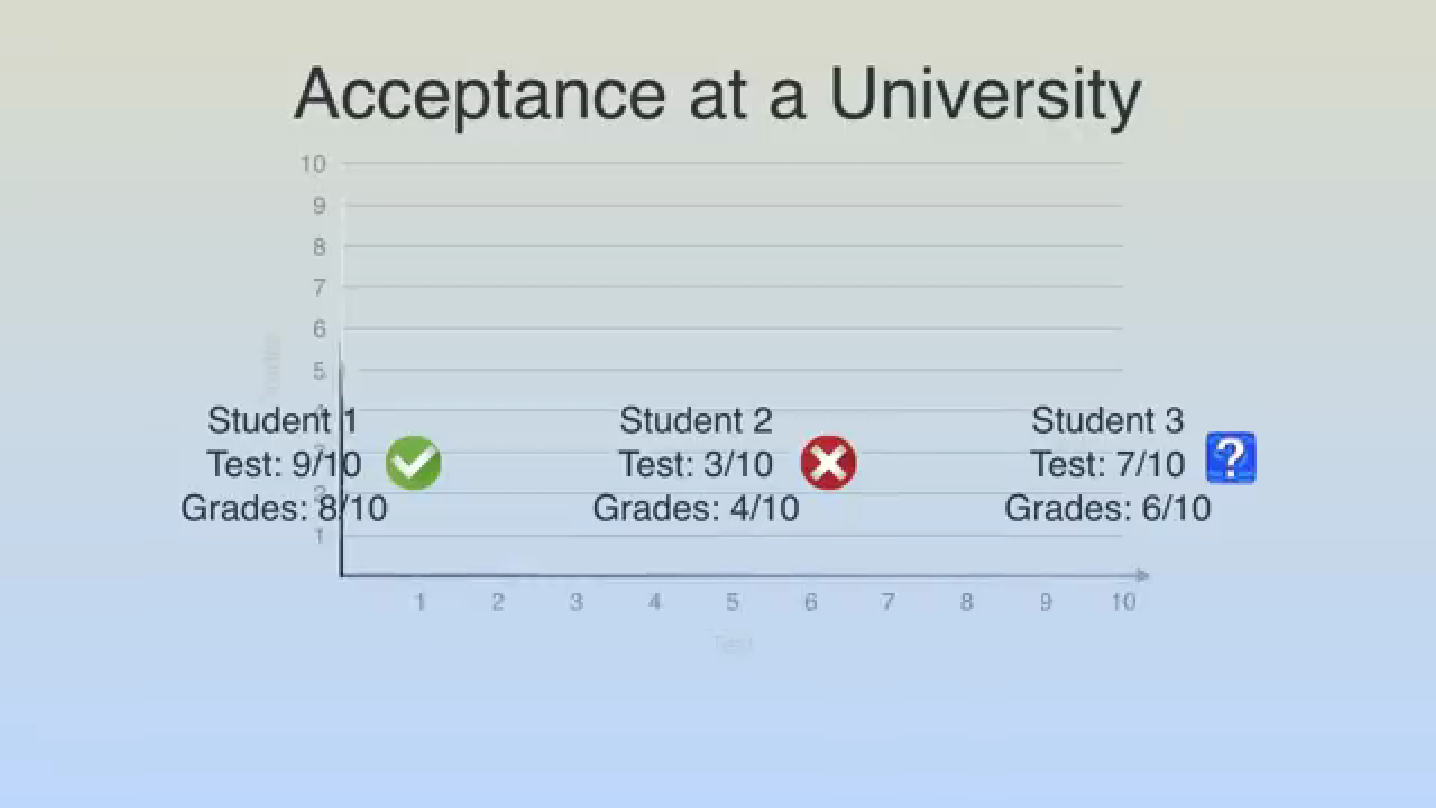


The text discusses an algorithm that aids in finding the best fitting data tree, using examples of app recommendations based on gender and university admissions decisions. The algorithm's goal is to optimize decision-making based on available data.

(137) A Friendly Introduction to Machine Learning - YouTube

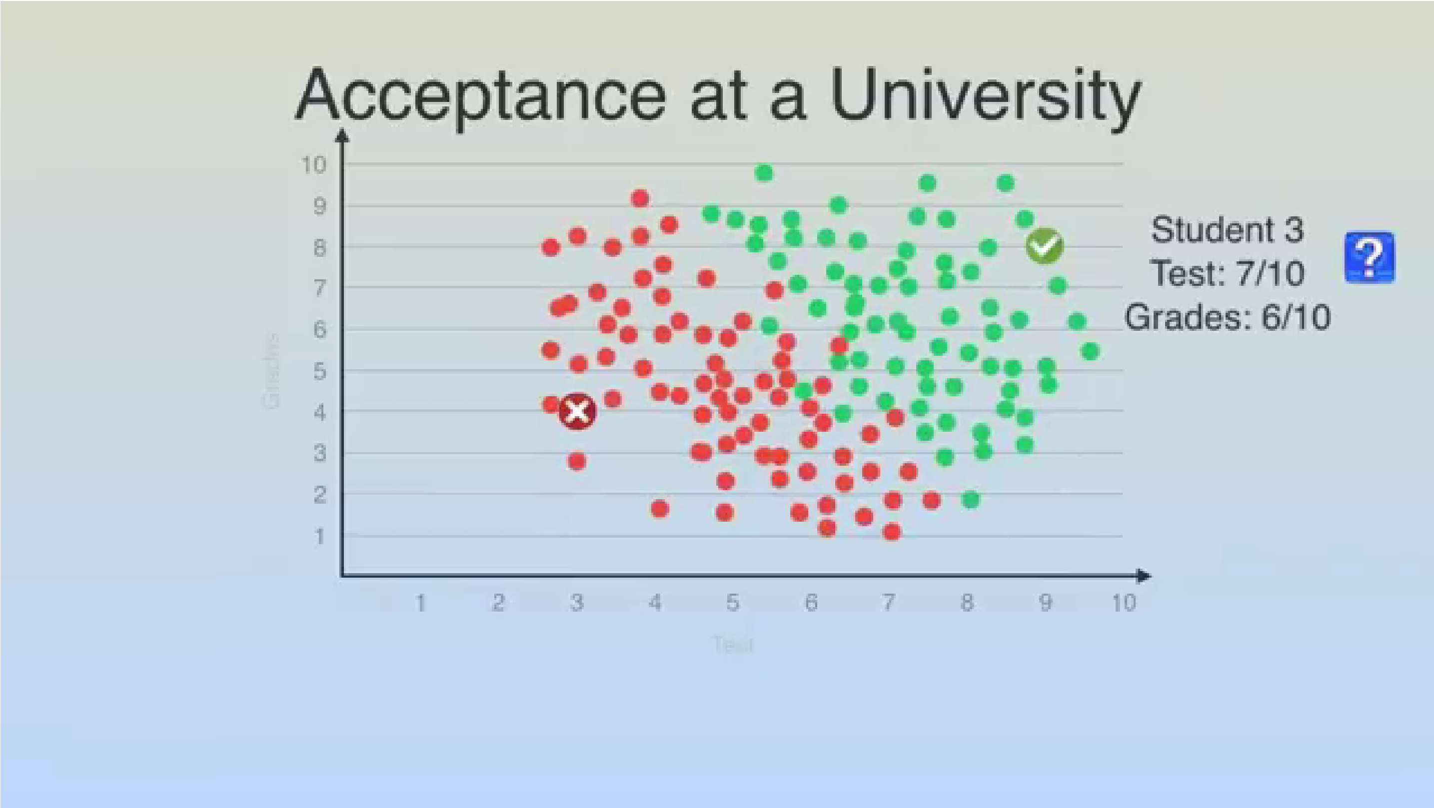
The entrance exam and school grades are the two criteria for student acceptance. For instance, Student One scored nine on the test and eight in grades, leading to acceptance, while Student Two's scores of three and four resulted in rejection. The acceptance of Student Three, with scores of seven and six, is

(137) A Friendly Introduction to Machine Learning - YouTube



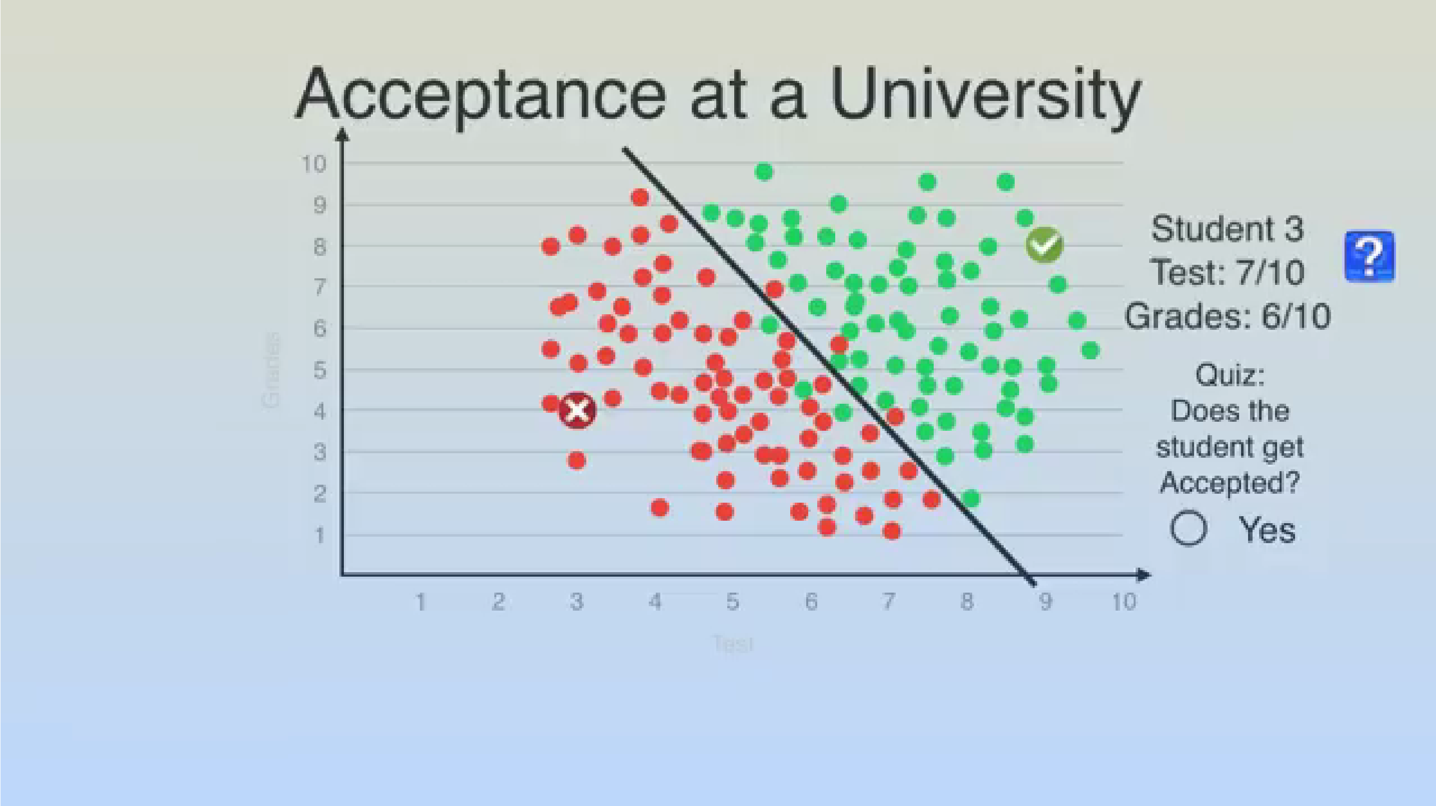
The x-axis and y-axis represent test scores and grades respectively. Student one's coordinates are 98, reflecting their scores of 9 and 8. Student two's coordinates are 34, indicating scores of 3 and 4. To decide on accepting or rejecting student three, a trend from

(137) A Friendly Introduction to Machine Learning - YouTube



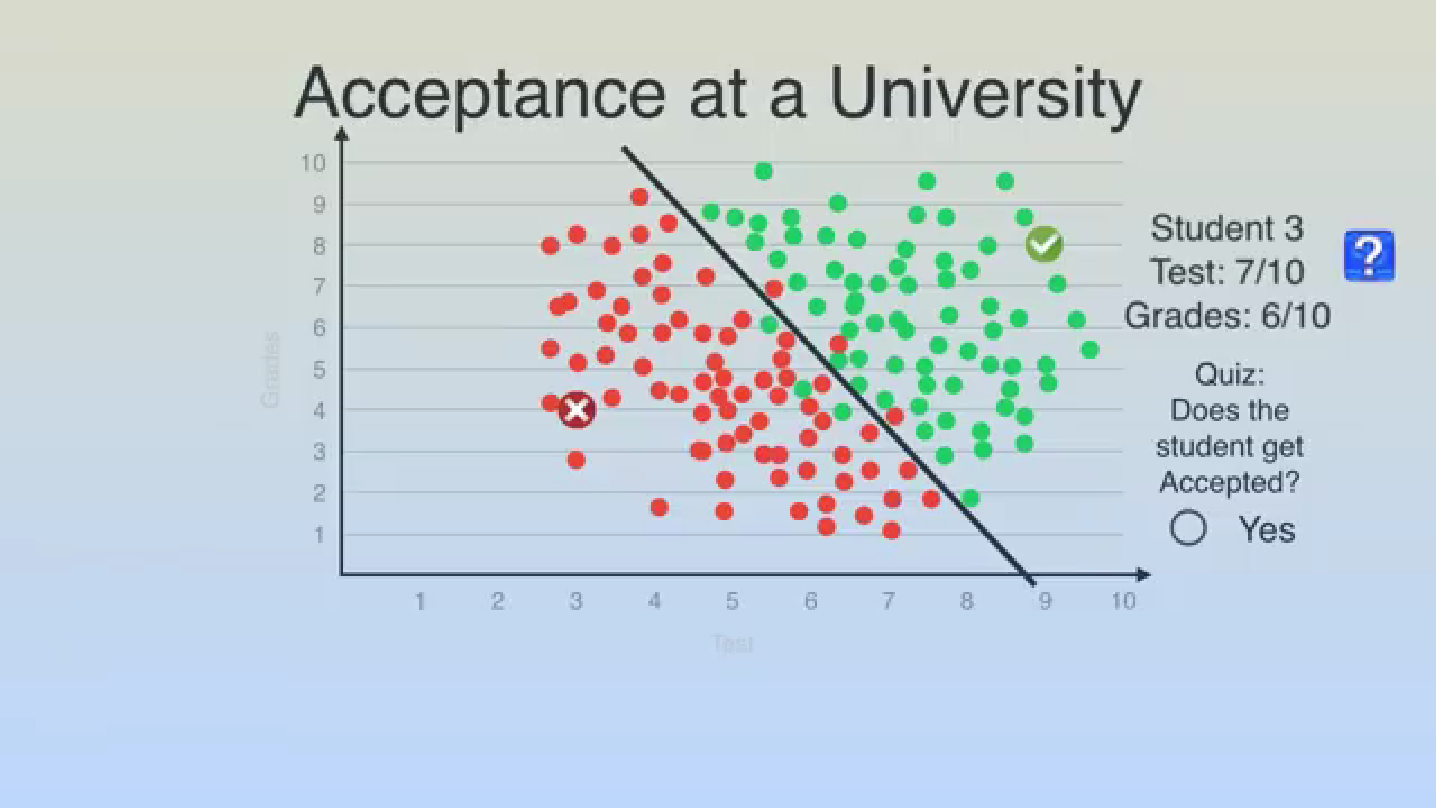
The green and red dots symbolize previously accepted and rejected students, respectively. Upon examining the data, there appears to be a line separating the dots, with mostly green dots above it and red ones below, albeit with some exceptions.

(137) A Friendly Introduction to Machine Learning - YouTube



The line in the graph represents a model used to predict student acceptance based on their scores. Students with high scores, above the line, are predicted to be accepted, while those with lower scores, below the line, are likely to be rejected.

(137) A Friendly Introduction to Machine Learning - YouTube

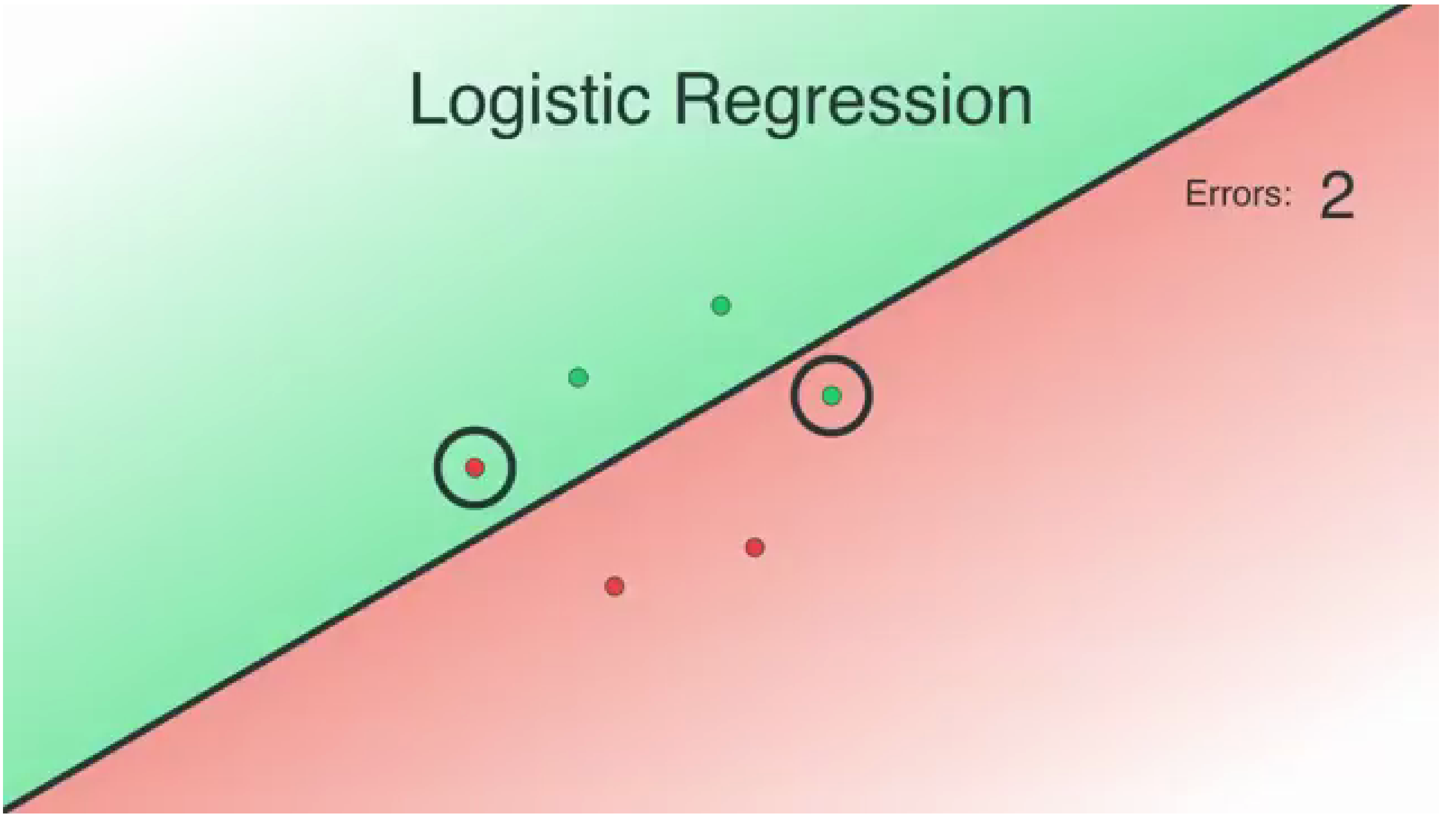


The individual is likely to reach point 76, thus being accepted, which is determined by logistic regression. This method involves finding a line that best separates data, exemplified by splitting six points, three green and three red, using computer analysis.

(137) A Friendly Introduction to Machine Learning - YouTube

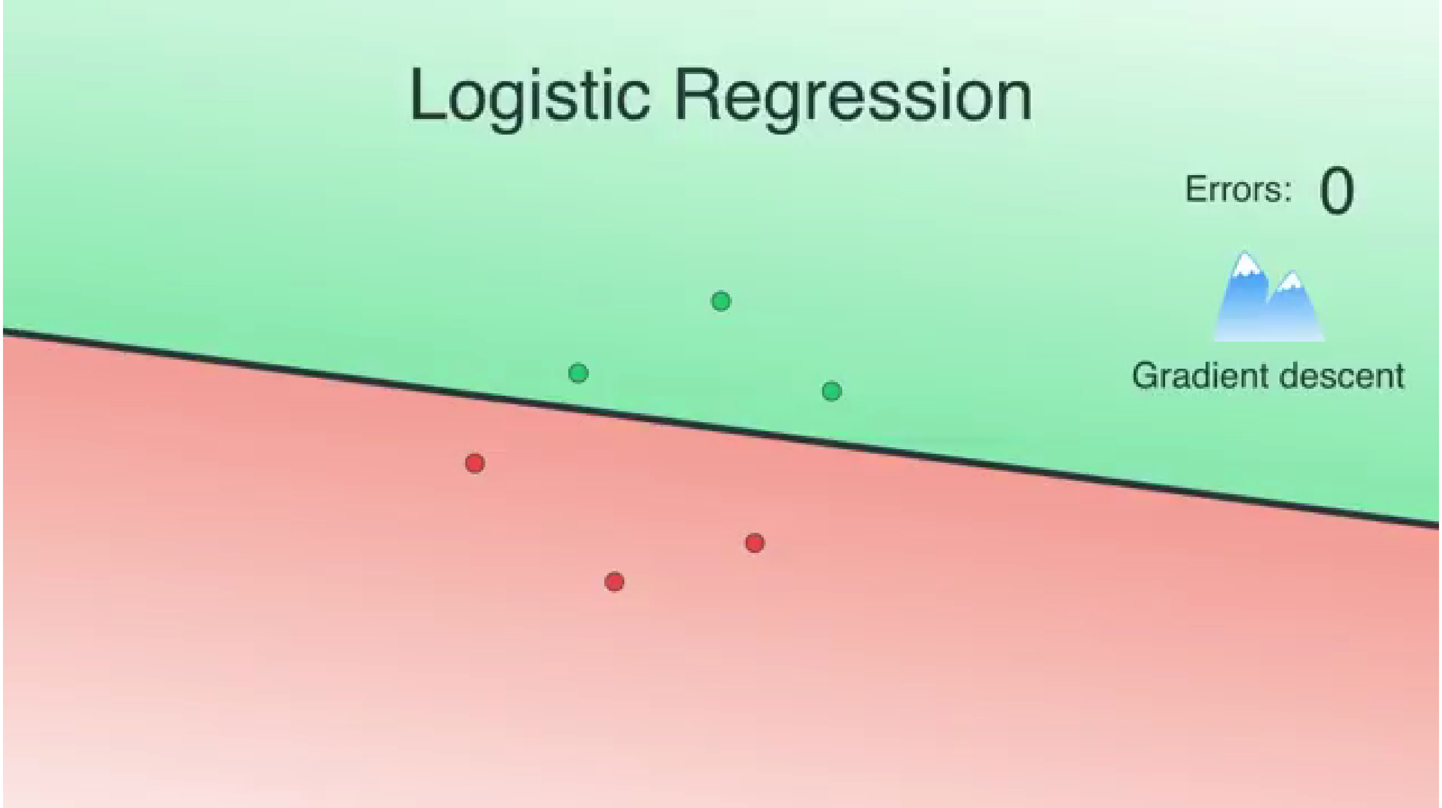
The lecture discusses drawing a random line and assigning regions as green or red, similar to linear regression. The effectiveness of the line is measured by the number of misclassified points, termed as 'error'. For instance, a line that misclassifies two points has an error of two.

(137) A Friendly Introduction to Machine Learning - YouTube



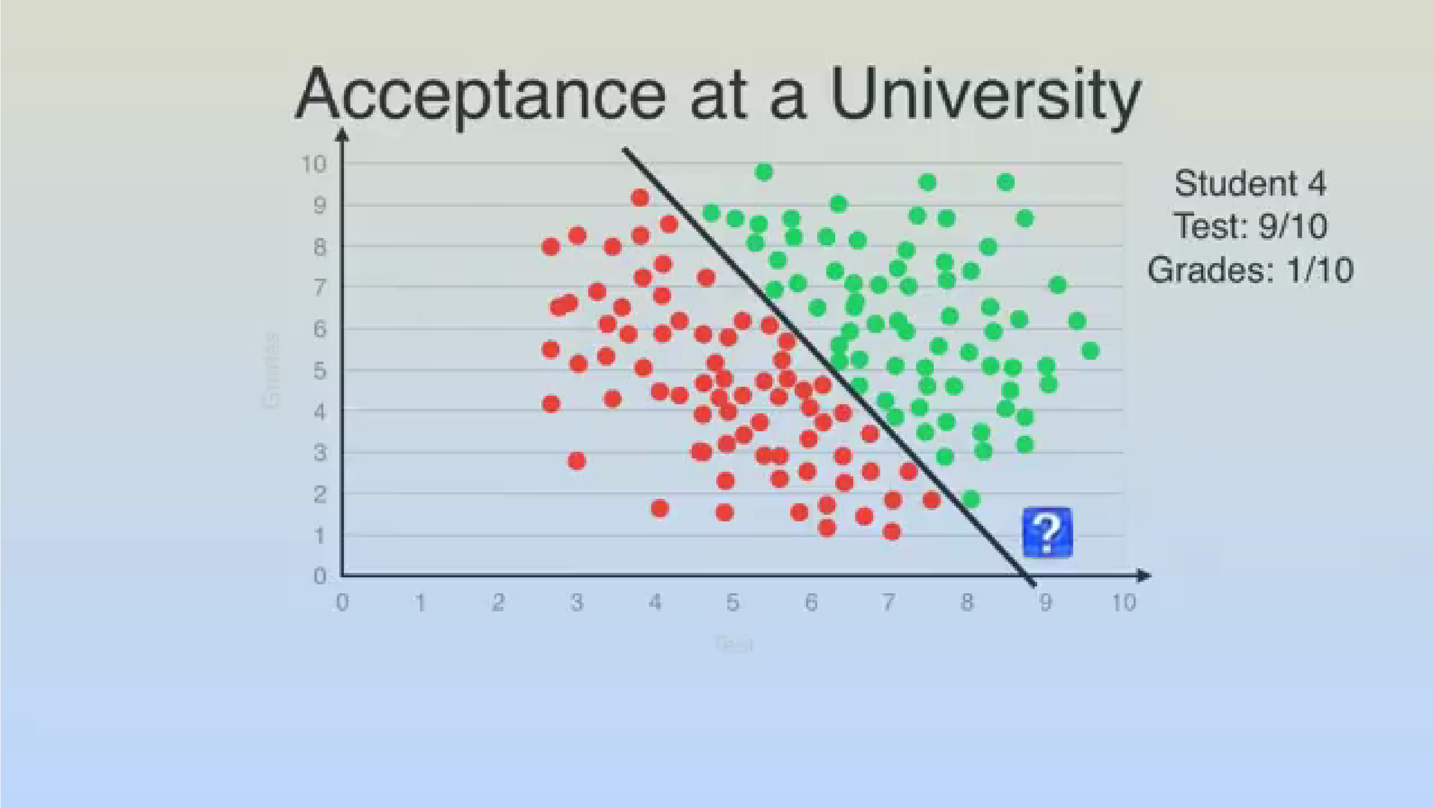
In linear regression, the line is adjusted to minimize errors using gradient descent. By shifting the line slightly, we can correctly classify data points, reducing the number of errors. Through further adjustments, we can potentially eliminate errors altogether. This process involves calculus and the gradient descent method.

(137) A Friendly Introduction to Machine Learning - YouTube



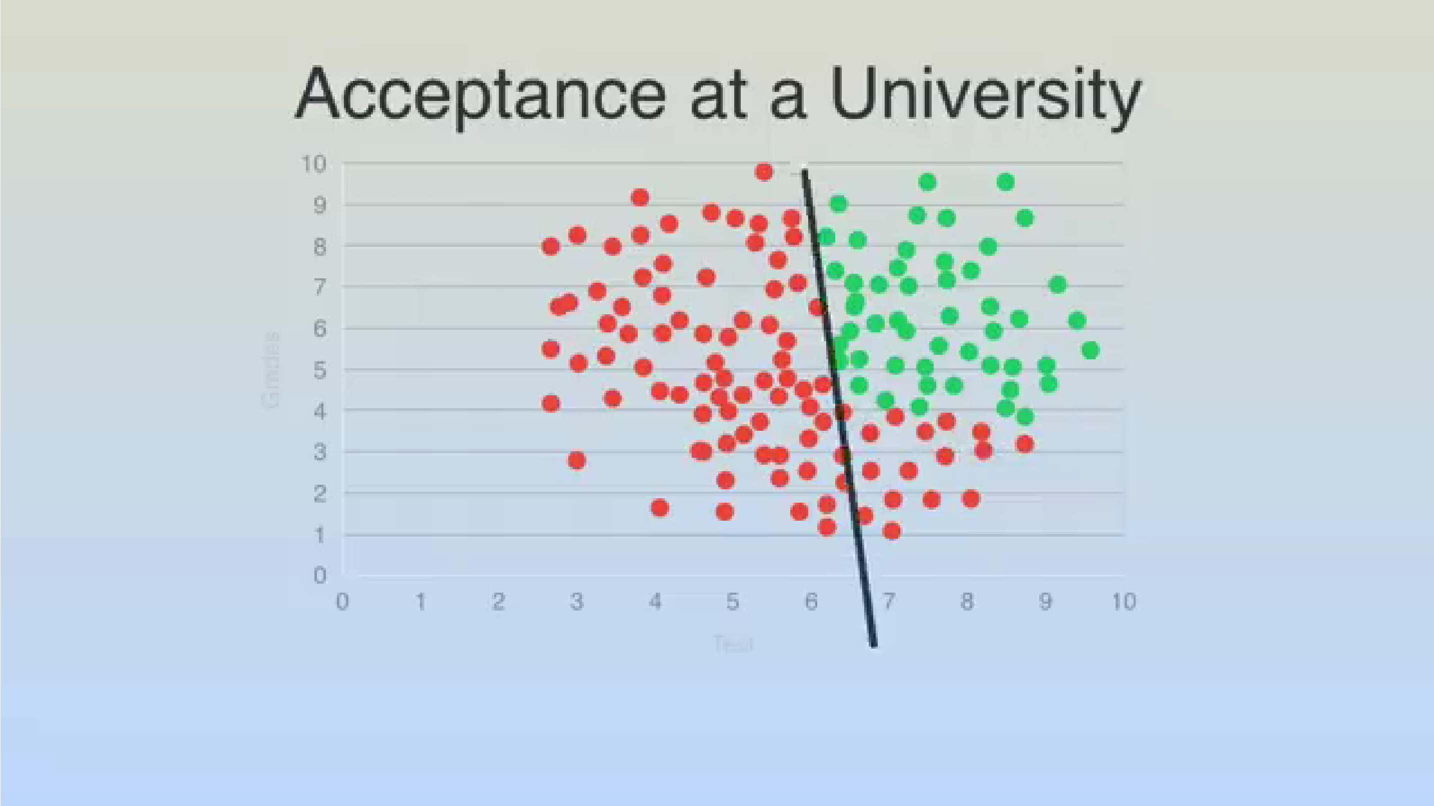
The Log loss function is a tool used to quantify the number of errors in a model, assigning a high value to misclassified points and a low value to correctly classified ones. For instance, in a model determining student acceptance, a student scoring nine on a test would be accepted according to our model.

(137) A Friendly Introduction to Machine Learning - YouTube



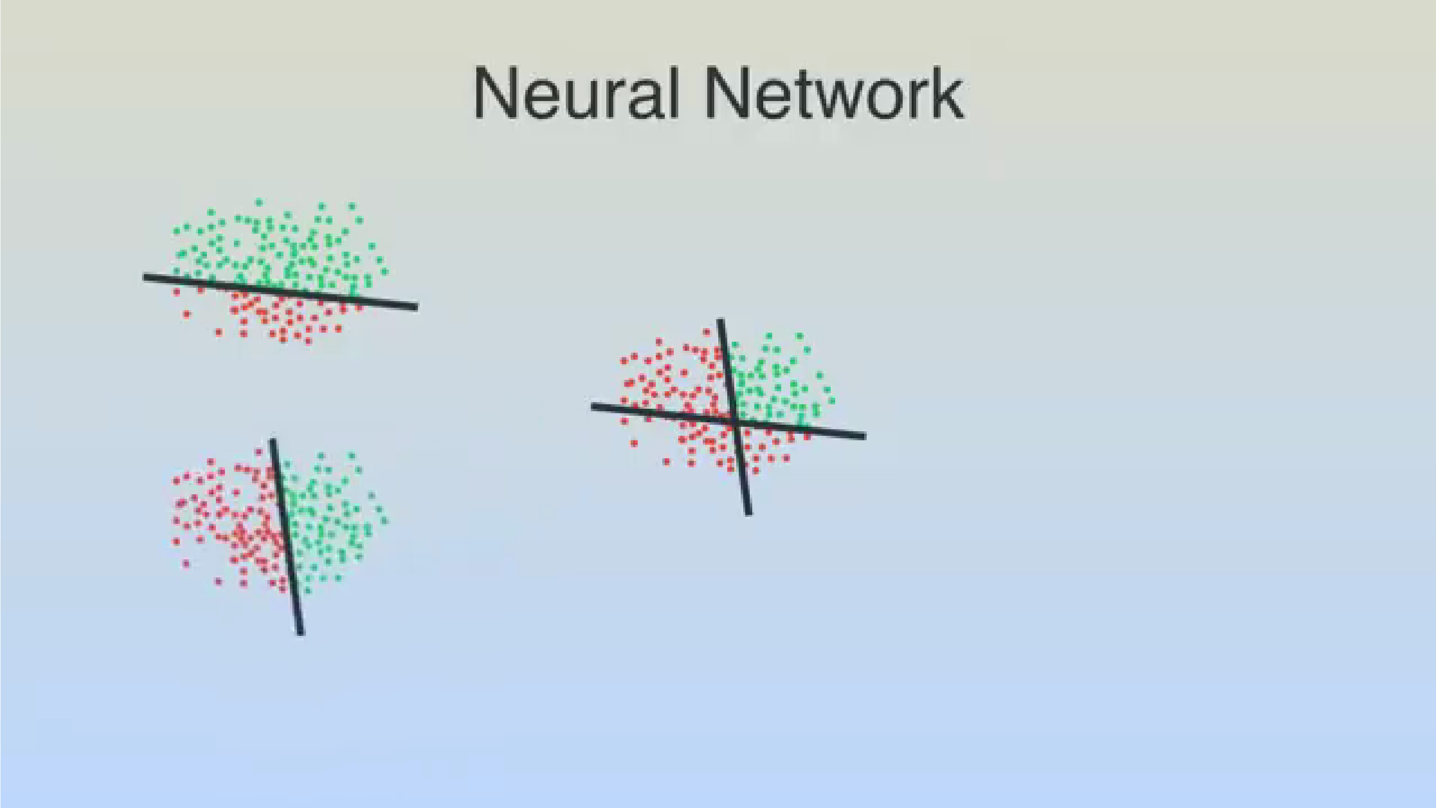
The acceptance of students should not be based solely on test scores, as this approach may overlook those with low grades. A single line division seems oversimplified. Instead, a more nuanced method, possibly involving a circular model or multiple lines, could be more effective.

(137) A Friendly Introduction to Machine Learning - YouTube



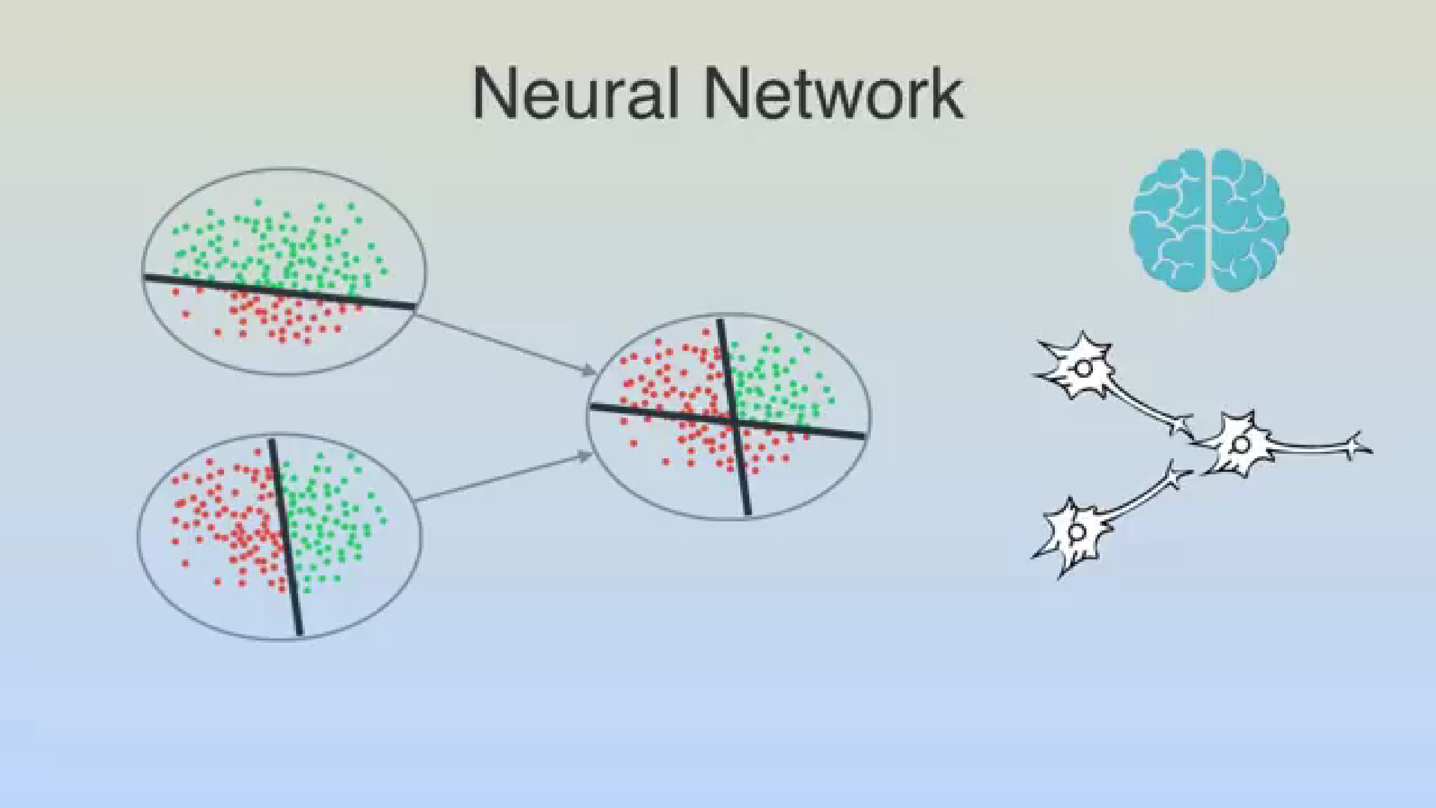
The optimal method involves using gradient descent to minimize a similar log loss function, which is referred to as a neural network. This term is derived from the intersection of two lines, creating a green area, which can be constructed as the intersection of the green area above one line and to the right of the

(137) A Friendly Introduction to Machine Learning - YouTube



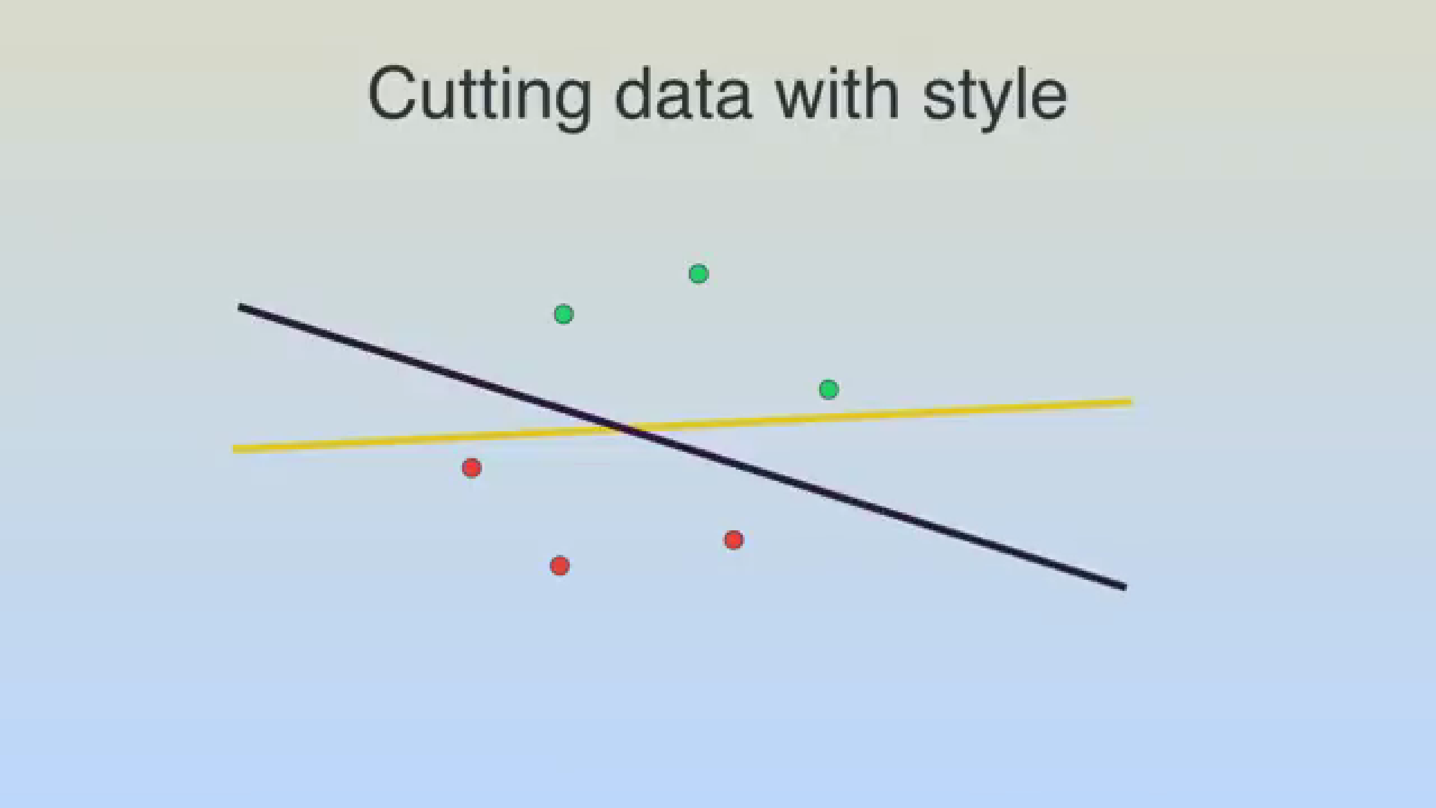
Graphing two nodes divides a plane into regions, with their intersection as the desired area, mimicking a neural network. This is similar to the brain's neurons that connect, fire electricity, and split the plane into regions based on whether a point belongs to one of those regions.

(137) A Friendly Introduction to Machine Learning - YouTube



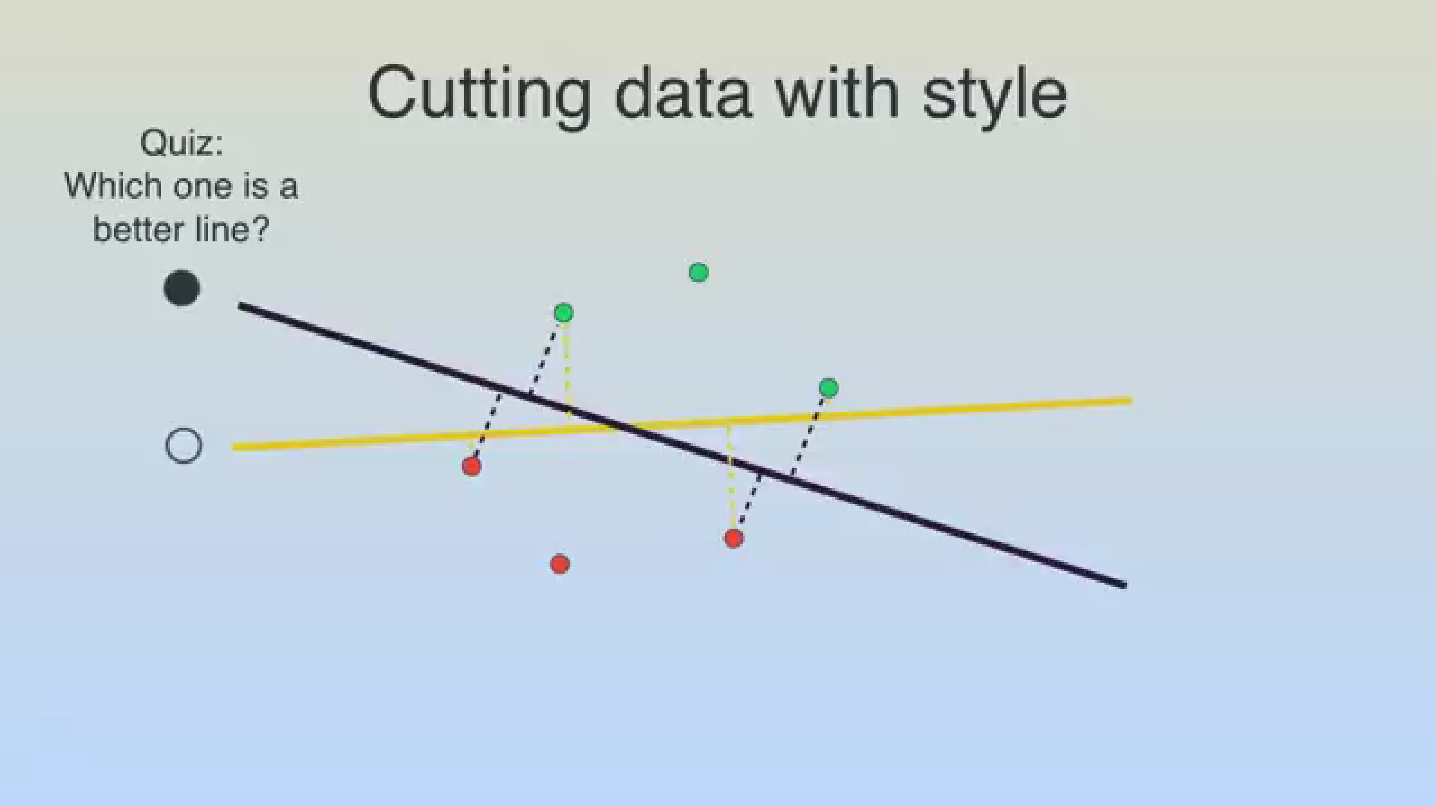
Linear regression can be viewed as a ninja, slicing data in half based on labels. Similarly, a neural network acts as a team of ninjas, dividing data into regions according to labels. When examining data points, there can be multiple lines, like yellow or purple, that can separate them.

(137) A Friendly Introduction to Machine Learning - YouTube



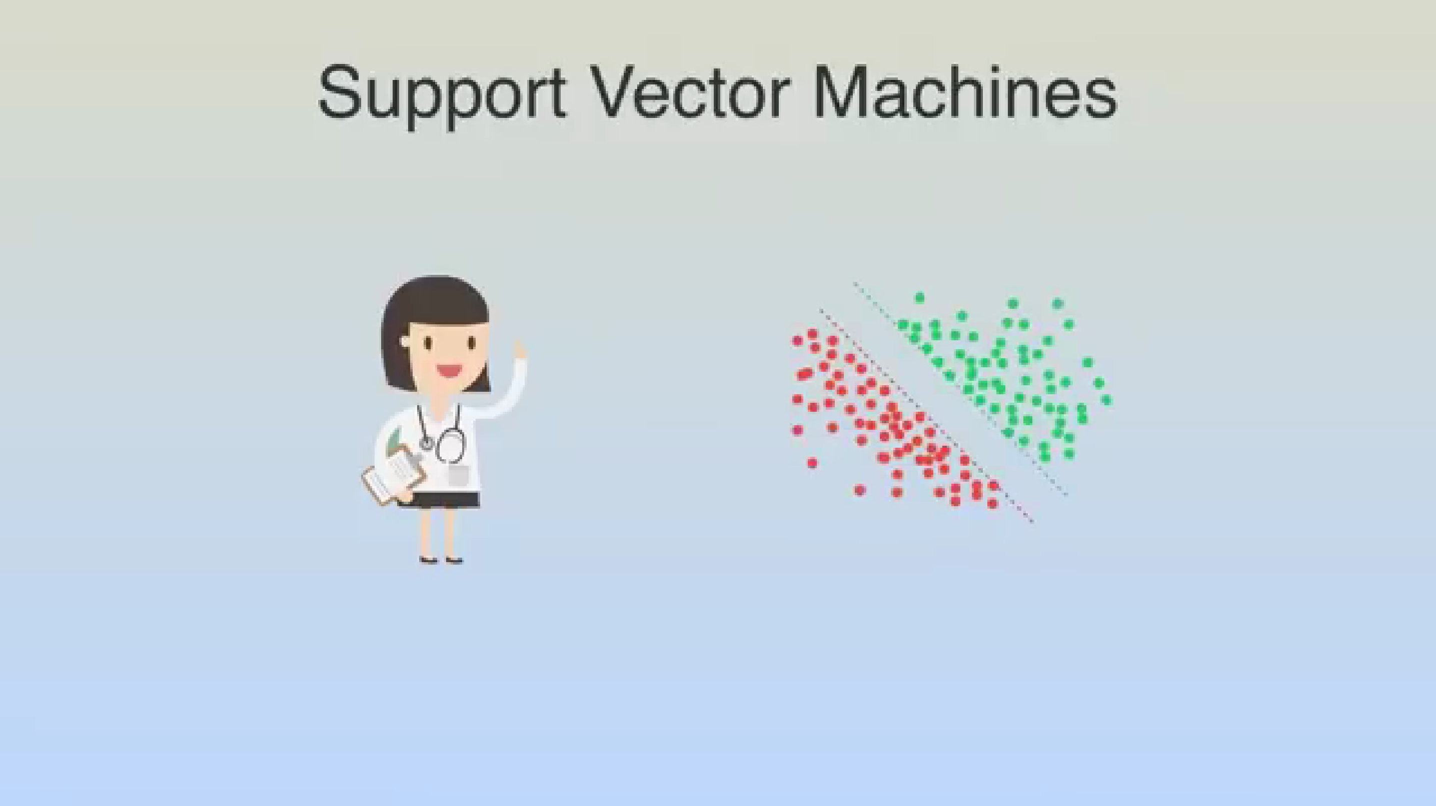
The purple line is considered the best for cutting the data as it is nicely spaced and far from all points, unlike the yellow line which is too close to two points and could misclassify some data. The challenge lies in determining how to find the purple line.

(137) A Friendly Introduction to Machine Learning - YouTube



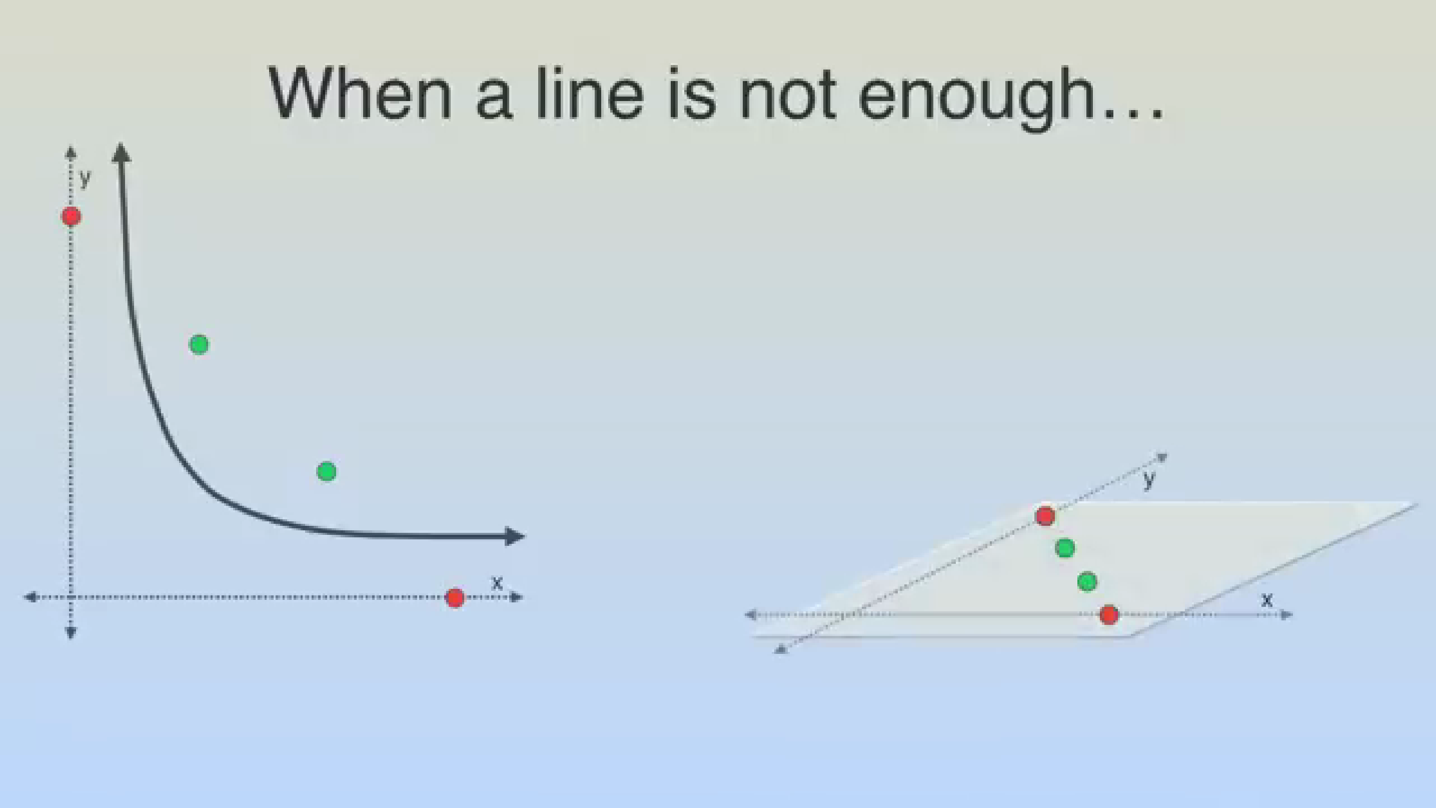
The key points are not to concern ourselves with points far from the boundary, but rather focus on those close by. Instead of using gradient descent, we'll apply linear optimization to find the line that maximizes distance from boundary points. This method is known as a Support Vector Machine, which can be liken

(137) A Friendly Introduction to Machine Learning - YouTube



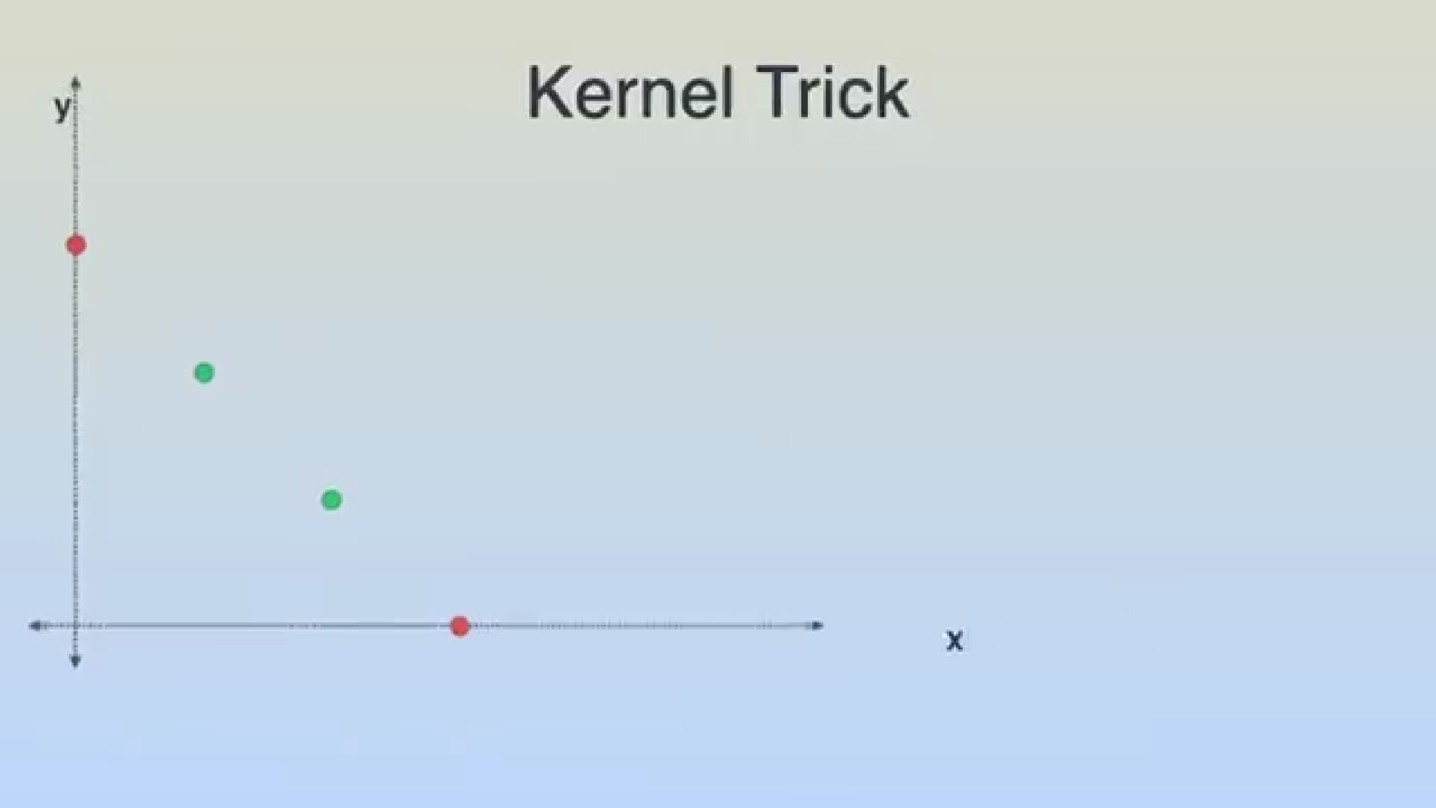
The data points can be divided using a line, but if they're distributed in a complex pattern, a curve might be more effective. Alternatively, considering the points within a three-dimensional space could offer another solution for separating them.

(137) A Friendly Introduction to Machine Learning - YouTube



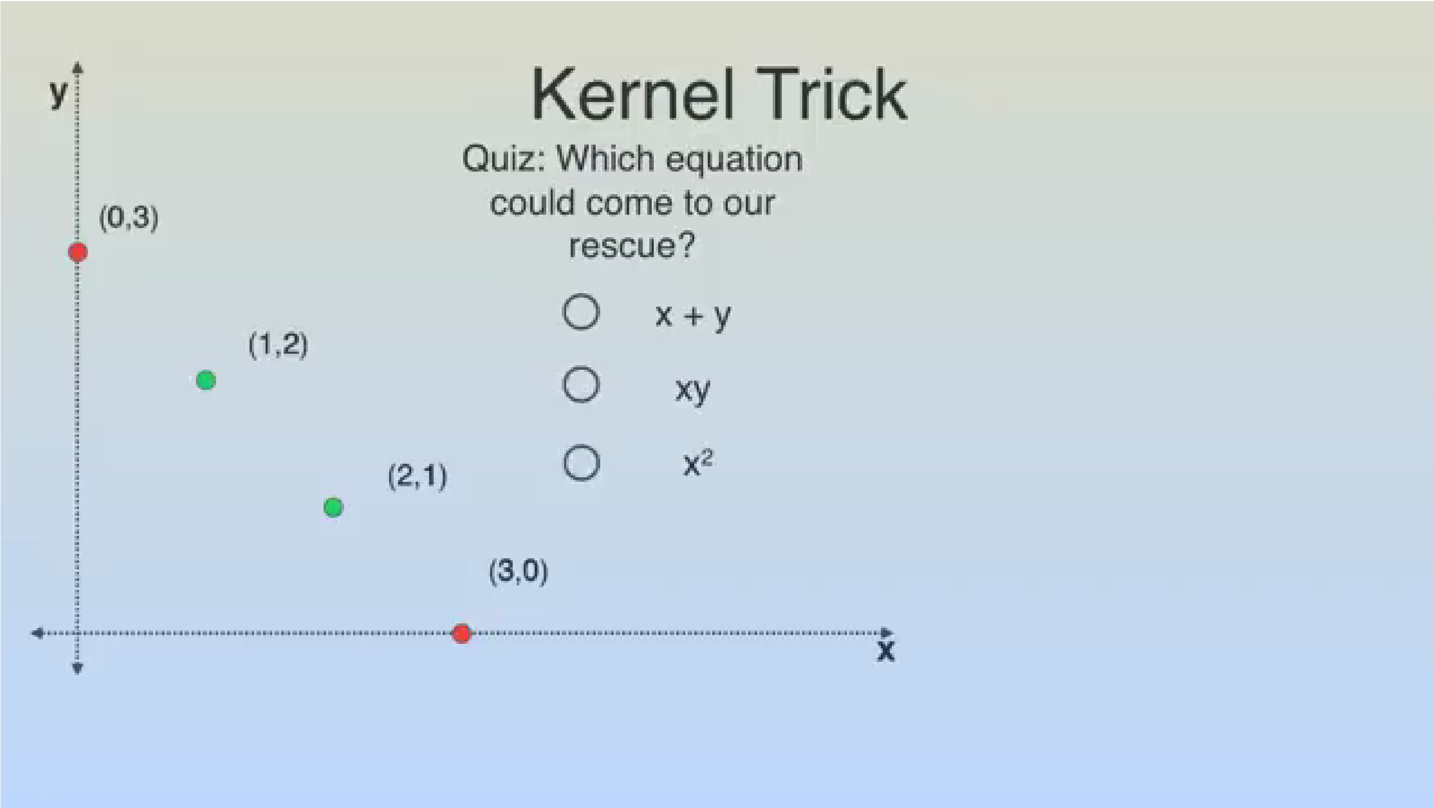
The lecture discusses the introduction of a third dimension, the z-axis, to separate two points on a plane. The solution involves a method called the kernel trick, commonly used in support vector machines. The seemingly different approaches of using a curve or a plane are actually the same. More details will be provided

(137) A Friendly Introduction to Machine Learning - YouTube



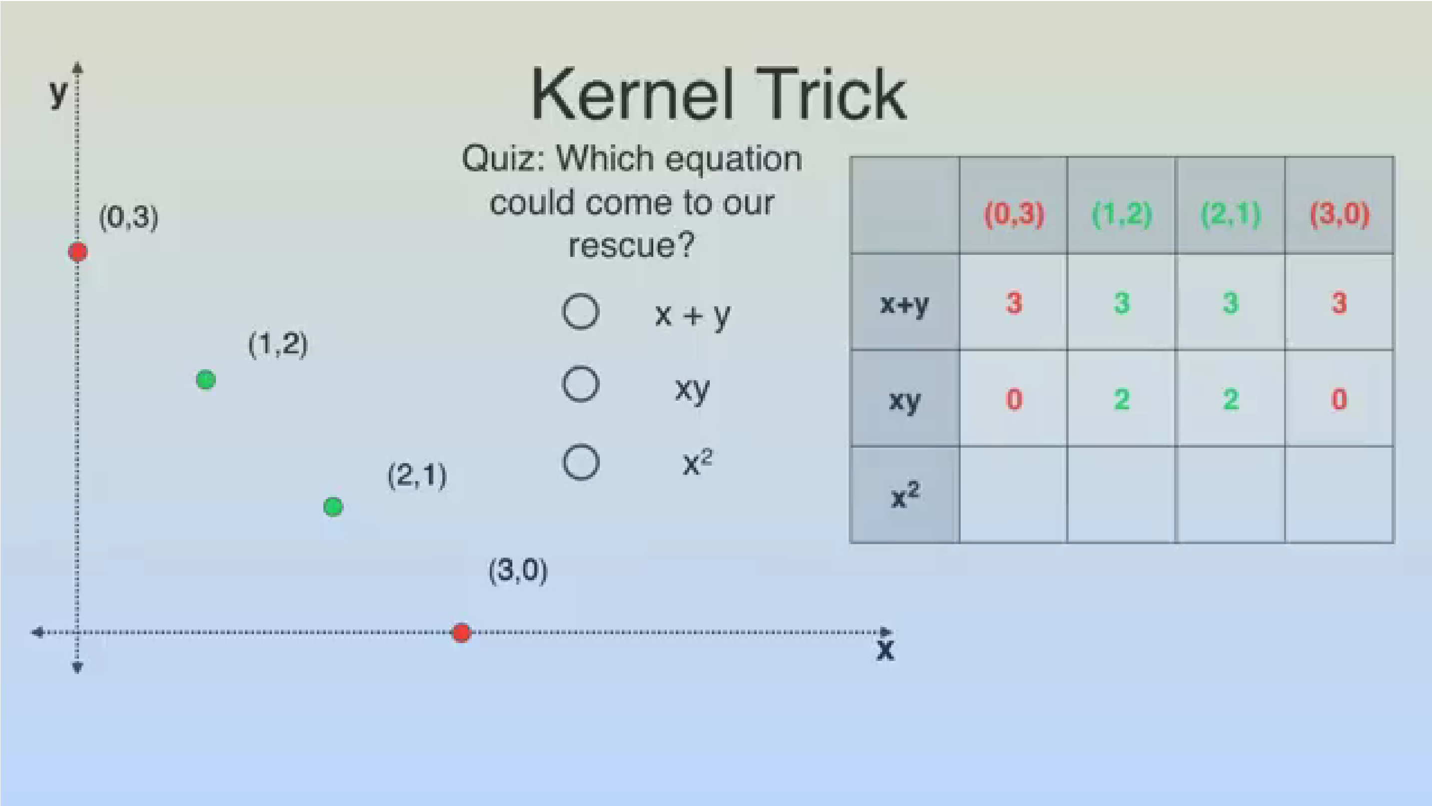
The lecture discusses the method of separating green points from red points on a coordinate plane. The points are assigned coordinates, and an equation involving these coordinates is needed to differentiate the points. The potential equations could be x + y or the product of x * y.

(137) A Friendly Introduction to Machine Learning - YouTube



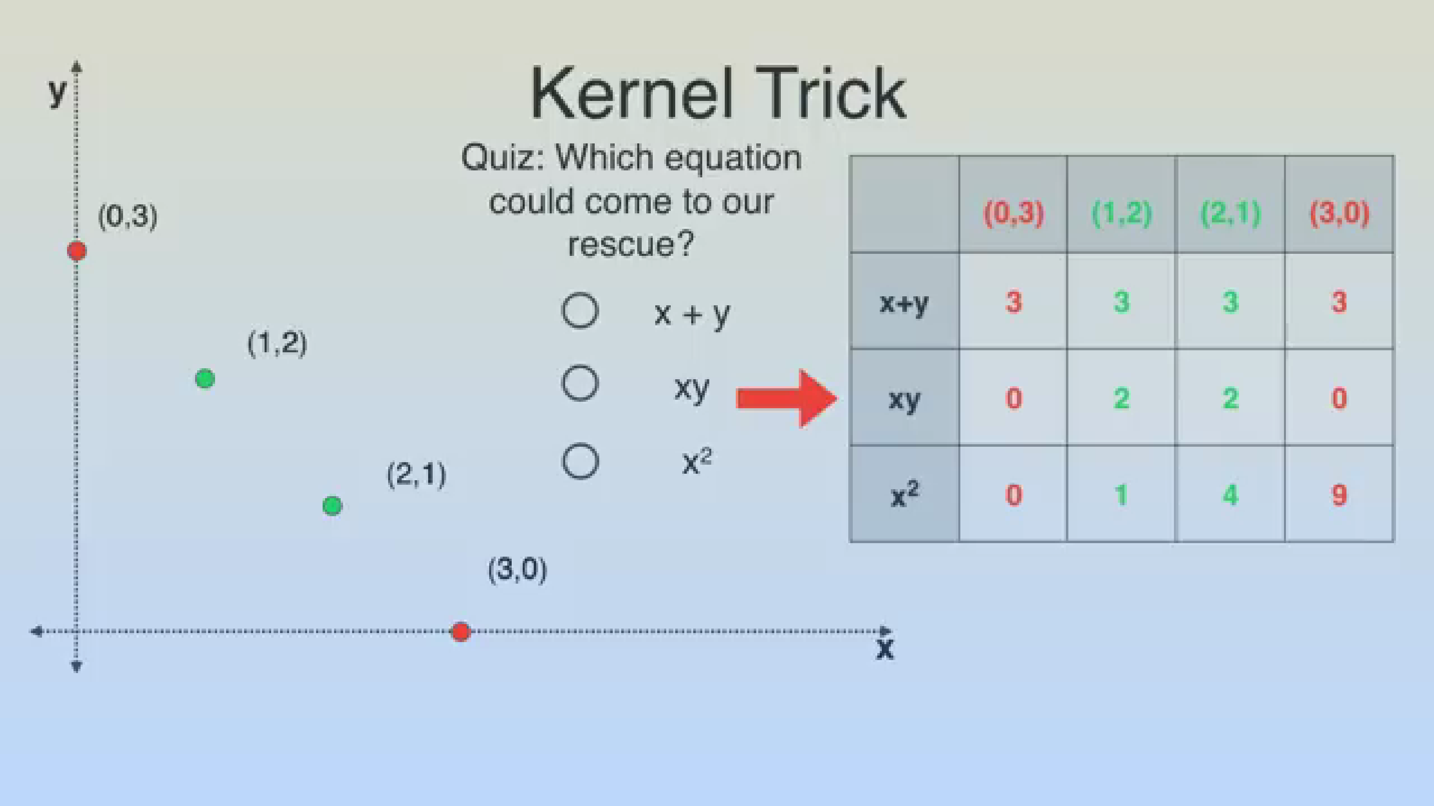
The text discusses creating a table to simplify a complex mathematical question. The table lists four points with their corresponding function values. The sum x+y and product x*y for each point pair are calculated, providing a clearer understanding of the equations.

(137) A Friendly Introduction to Machine Learning - YouTube



The lecture discusses the difficulty of distinguishing between red and green points using equations. The sum of x and y doesn't differentiate them, nor does the square of x. However, the product of x and y yields zero for the red points.

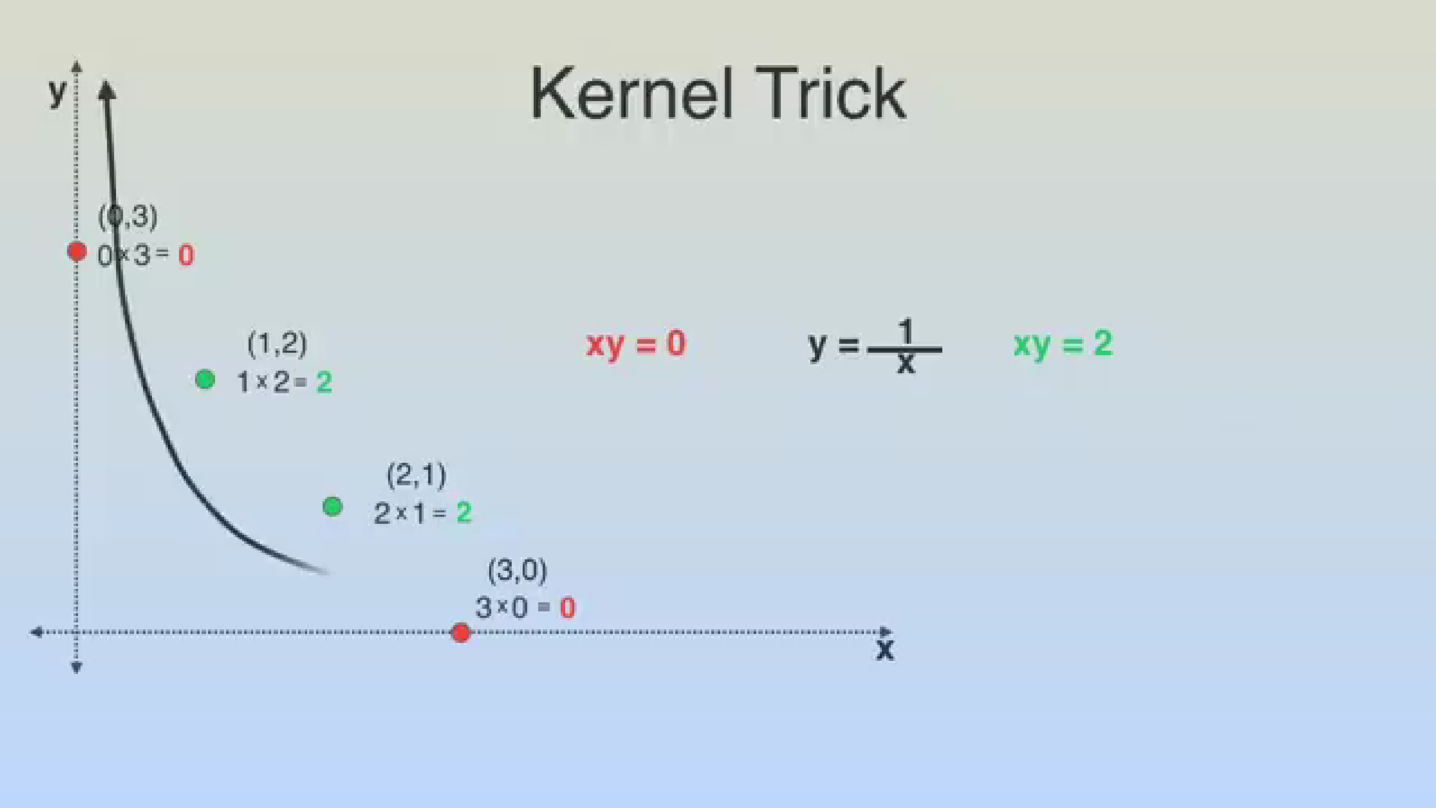
(137) A Friendly Introduction to Machine Learning - YouTube



	(0,3)	(1,2)	(2,1)	(3,0)
x+y	3	3	3	3
xy	0	2	2	0
x^2	0	1	4	9

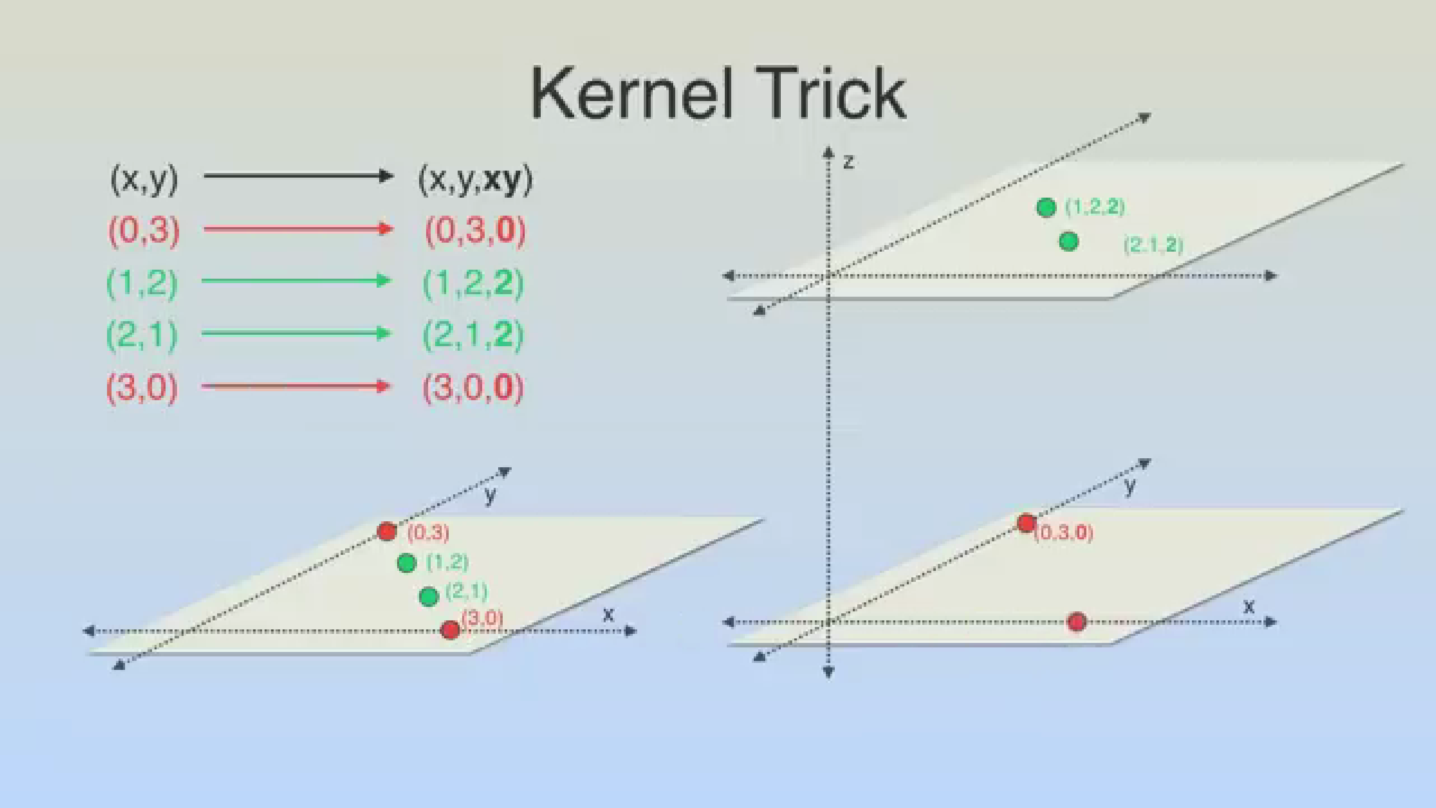
The function that separates red and green points is represented by the equation x*y=1. This equation is equivalent to y=1/x, which graphically corresponds to a hyperbola. For red points, the product x*y equals Z, while for green points, the product x*y equals

(137) A Friendly Introduction to Machine Learning - YouTube



The kernel trick involves mapping points from a curve to three-dimensional space. Points are mapped from (x, y) to (x, y, x*y), where the third coordinate represents height. For example, point (0,3) maps to (0,3,0), (1,

(137) A Friendly Introduction to Machine Learning - YouTube



The support vector machine kernel method is likened to a surgeon trying to separate apples and oranges. The solution involves moving the apples up and the oranges down, then drawing a line between them. The example shifts to a scenario involving the strategic placement of three pizza parlors in a city based on a study

(137) A Friendly Introduction to Machine Learning - YouTube

The text discusses the strategic placement of three pizza parlors based on the division of houses into three clusters: red, blue, and yellow. It emphasizes the need for an algorithm, rather than mere observation, to teach a computer to determine these optimal locations.

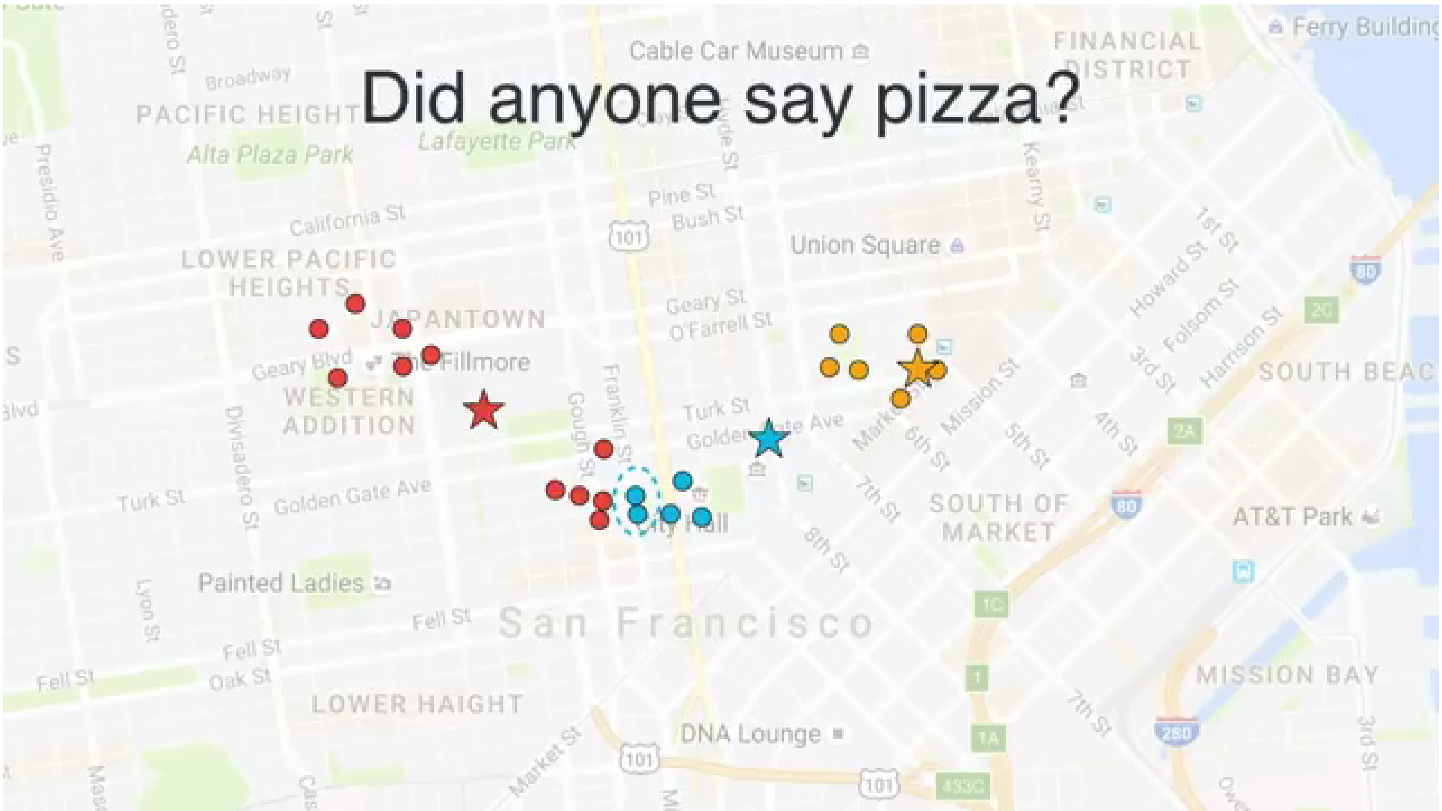
(137) A Friendly Introduction to Machine Learning - YouTube

Pizza parlors are randomly located at red, blue, and yellow stars. Each house should ideally go to the nearest pizza parlor. However, considering the location of the yellow houses, it would be more logical to relocate the yellow pizza parlor to their center.

(137) A Friendly Introduction to Machine Learning - YouTube

The lecture discusses repositioning pizza parlors to the center of the houses they serve. It suggests recoloring blue points closer to the yellow pizza parlor as yellow, and red points closer to the blue parlor as blue, for optimal service.

(137) A Friendly Introduction to Machine Learning - YouTube



The K-means clustering algorithm aids in determining the optimal locations for pizza parlors. By starting with random points and iterating the process, the best locations are achieved. However, there may be situations where the number of clusters isn't specified initially.

(137) A Friendly Introduction to Machine Learning - YouTube

The text suggests a method for grouping houses based on their proximity for pizza delivery. The closest houses are grouped together, assuming they will be served by the same pizza parlor. This process is repeated for the next nearest houses and so on.

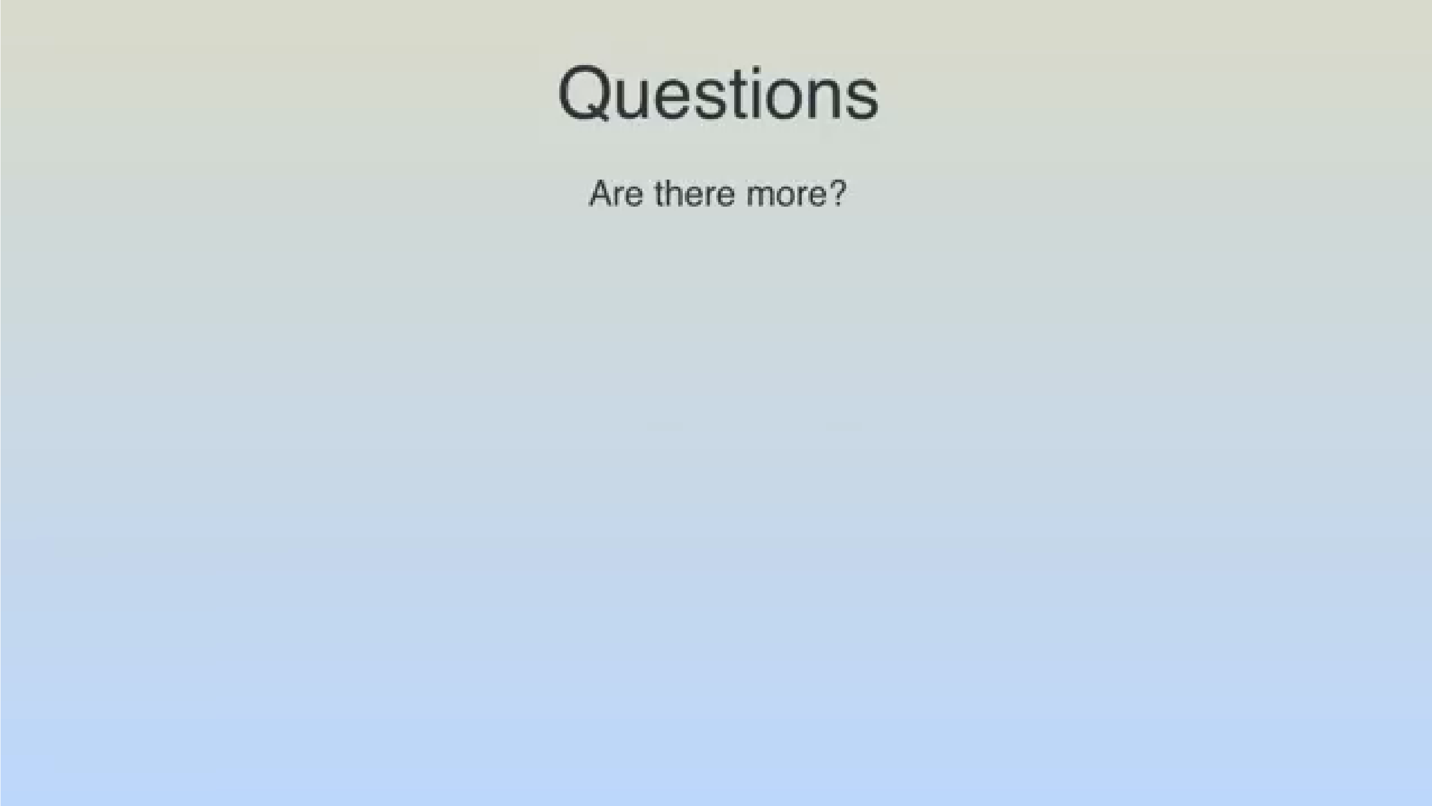
(137) A Friendly Introduction to Machine Learning - YouTube

The lecture covers hierarchical clustering, a main algorithm in machine learning. It involves grouping houses based on proximity, with a maximum distance specified to prevent overly large clusters. This process continues until the desired clusters are achieved.

(137) A Friendly Introduction to Machine Learning - YouTube

The lecture covered various algorithms and their applications, such as using linear regression for pricing, naive bayes for spam detection, decision trees for app recommendations, logistic regression for admission office models, and neural networks and support vector machines for improvement. It also discussed using clustering algorithms for locating pizza parlors. The

(137) A Friendly Introduction to Machine Learning - YouTube



Choosing the right algorithm for a given data set can be challenging. It's crucial to compare, evaluate, and understand which algorithm performs better in terms of running time and accuracy. The Udacity machine learning Nanodegree provides answers to these concerns and offers practical examples and projects.