Machine Learning 101

Basics

Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

1. Definition of Machine Learning

Machine Learning (ML) is a subset of artificial intelligence that focuses on developing algorithms and statistical models that enable computer systems to improve their performance on a specific task through experience, without being explicitly programmed.

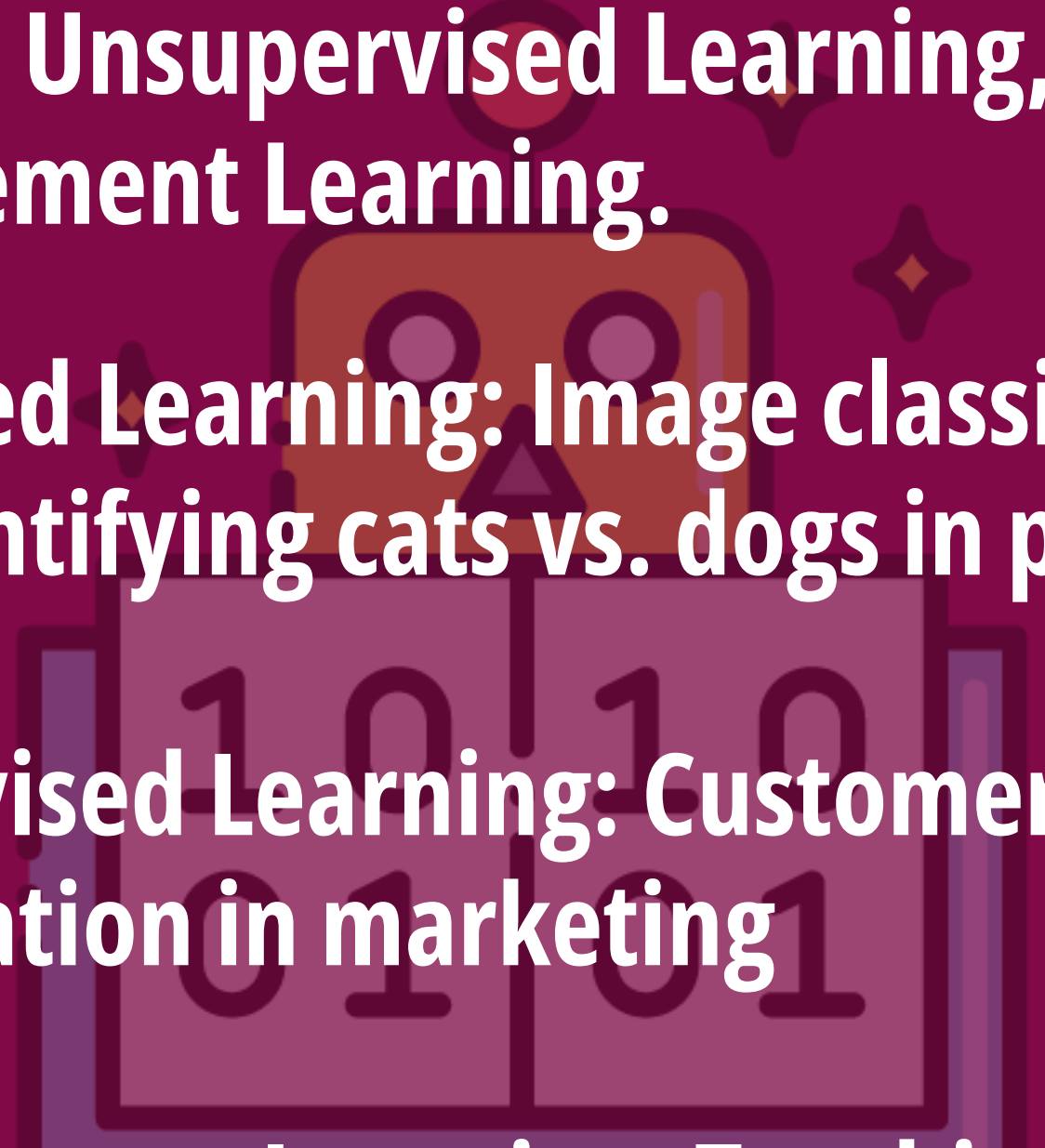
An email spam filter that learns to distinguish between spam and legitimate emails based on patterns it observes in large datasets of pre-classified emails.

Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

2. Types of Machine Learning

There are three main types: Supervised Learning, Unsupervised Learning, and Reinforcement Learning.



Supervised Learning: Image classification (e.g., identifying cats vs. dogs in pictures)

Unsupervised Learning: Customer segmentation in marketing

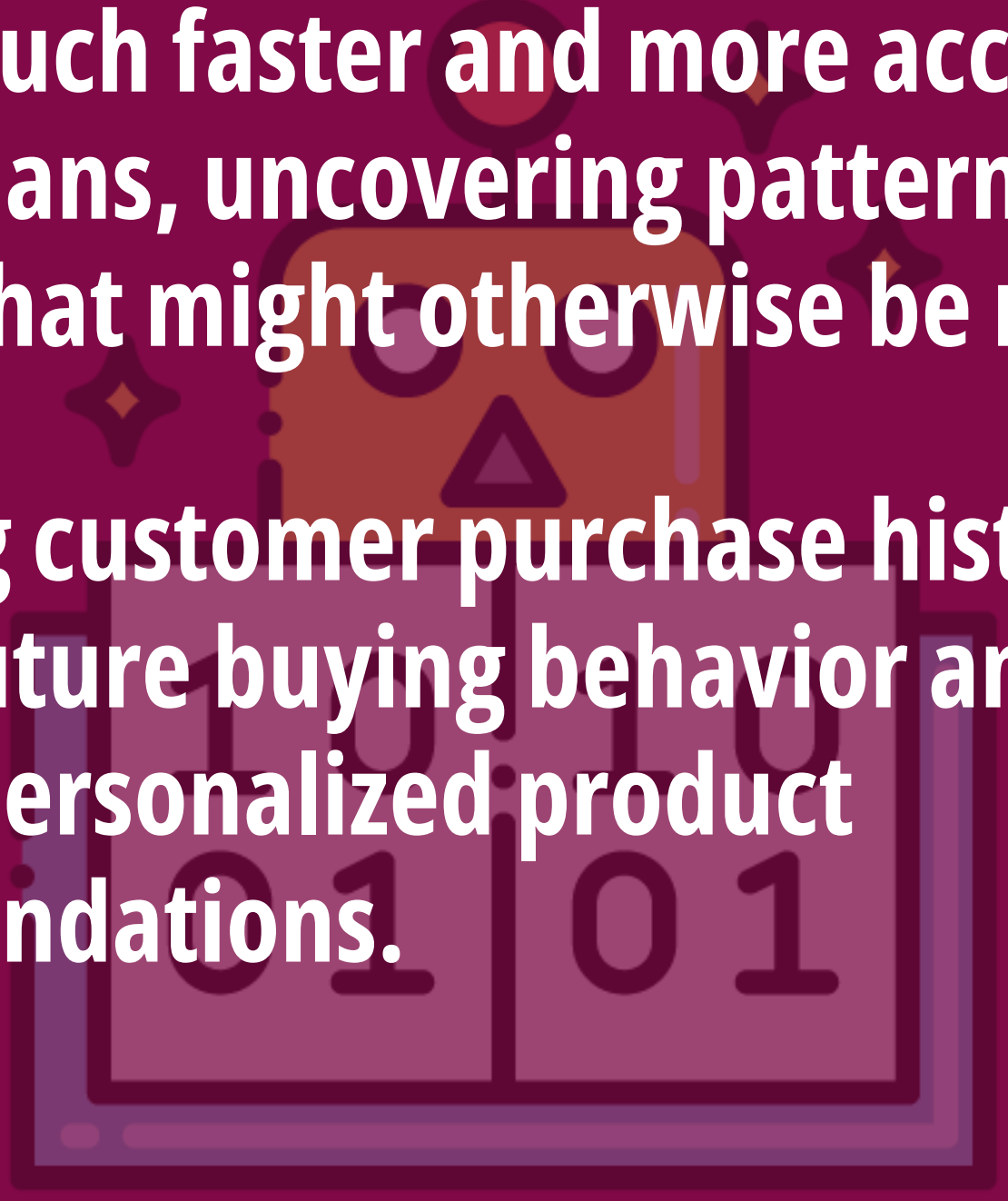
Reinforcement Learning: Teaching a computer to play chess

Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

3. Importance in Data Analysis

ML can process and analyze vast amounts of data much faster and more accurately than humans, uncovering patterns and insights that might otherwise be missed.



Analyzing customer purchase history to predict future buying behavior and provide personalized product recommendations.

Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

4. Automation and Efficiency

ML enables the automation of complex tasks, increasing efficiency and reducing human error in various industries.

Autonomous vehicles using ML algorithms to navigate roads and make real-time decisions.

Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

5. Predictive Capabilities

ML models can make accurate predictions based on historical data, helping businesses and organizations make informed decisions.

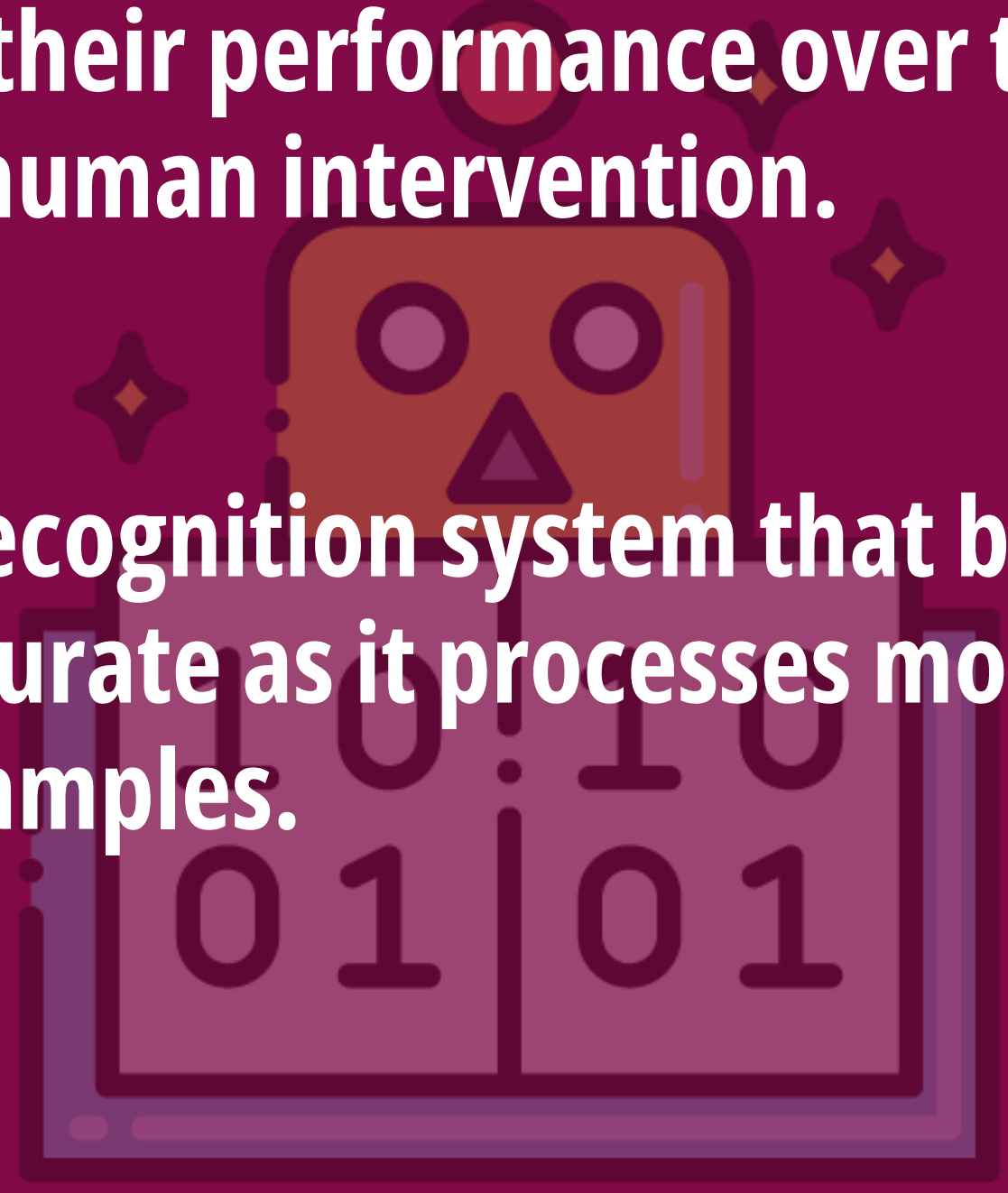
Predicting stock prices based on historical market data and current economic indicators.

Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

6. Adaptability

ML systems can adapt to new data and improve their performance over time without human intervention.



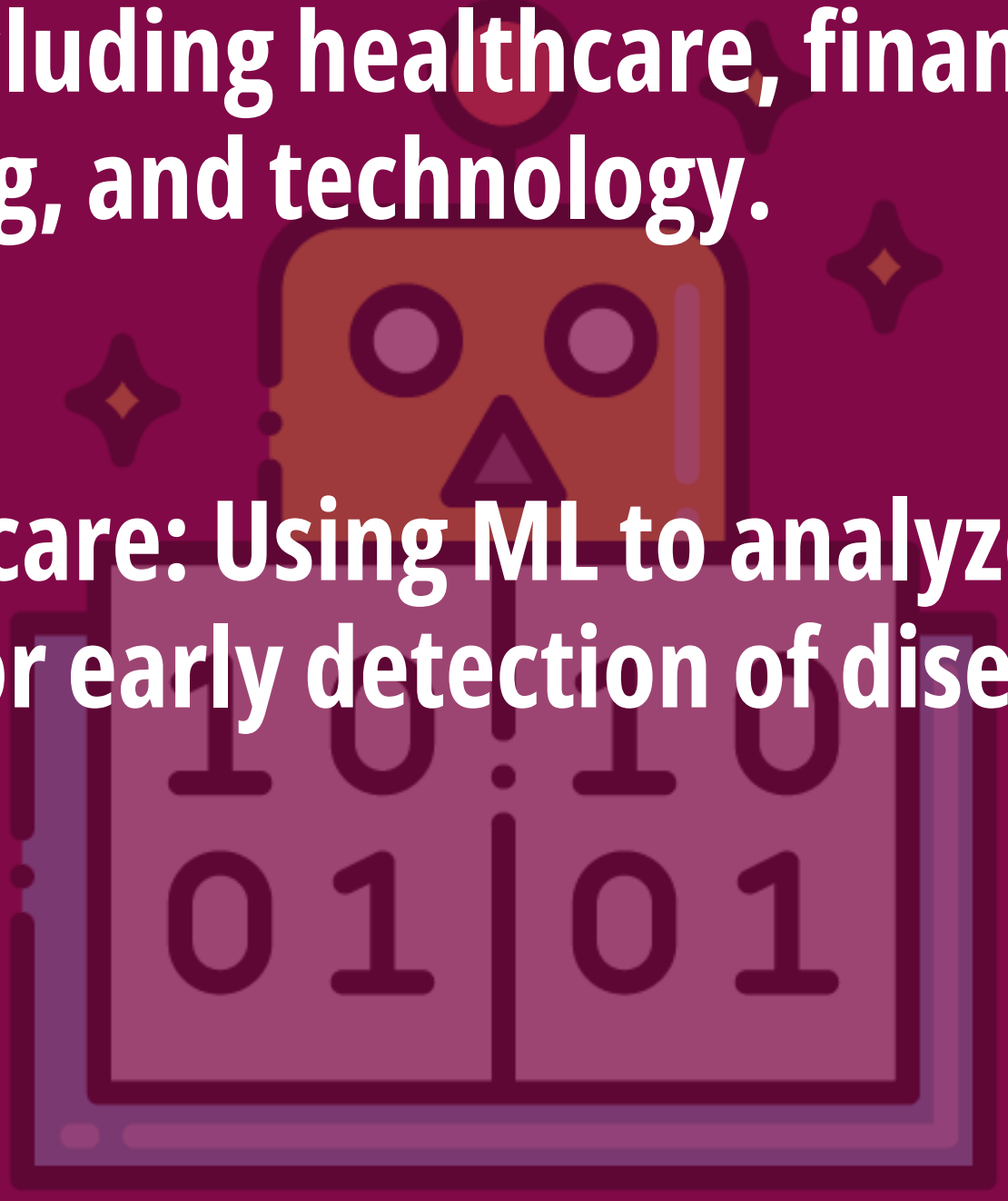
A voice recognition system that becomes more accurate as it processes more diverse speech samples.

Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

7. Wide Range of Applications

ML has applications across numerous fields, including healthcare, finance, marketing, and technology.



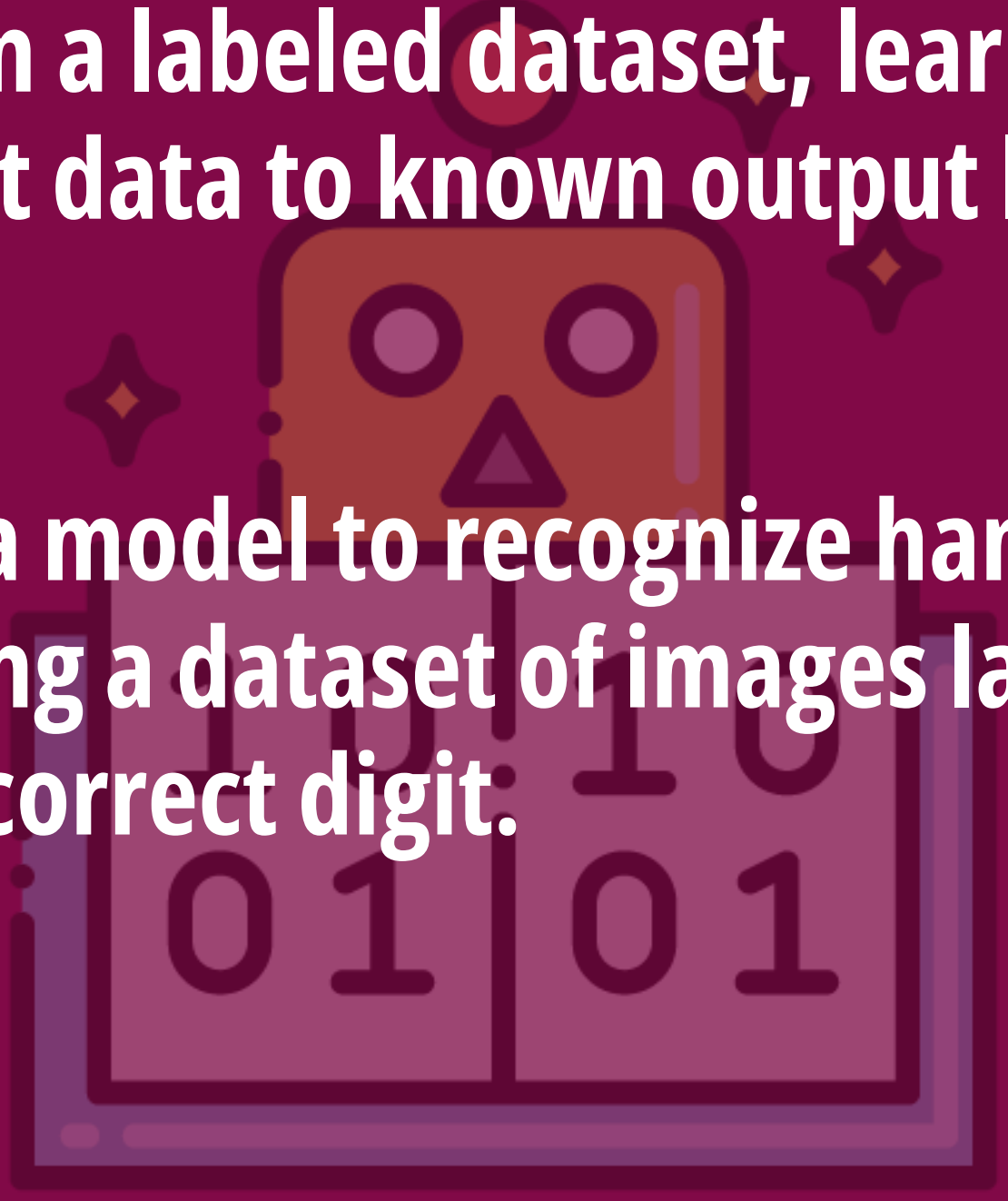
In healthcare: Using ML to analyze medical images for early detection of diseases like cancer.

Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

8. Supervised Learning

A type of ML where the algorithm is trained on a labeled dataset, learning to map input data to known output labels.



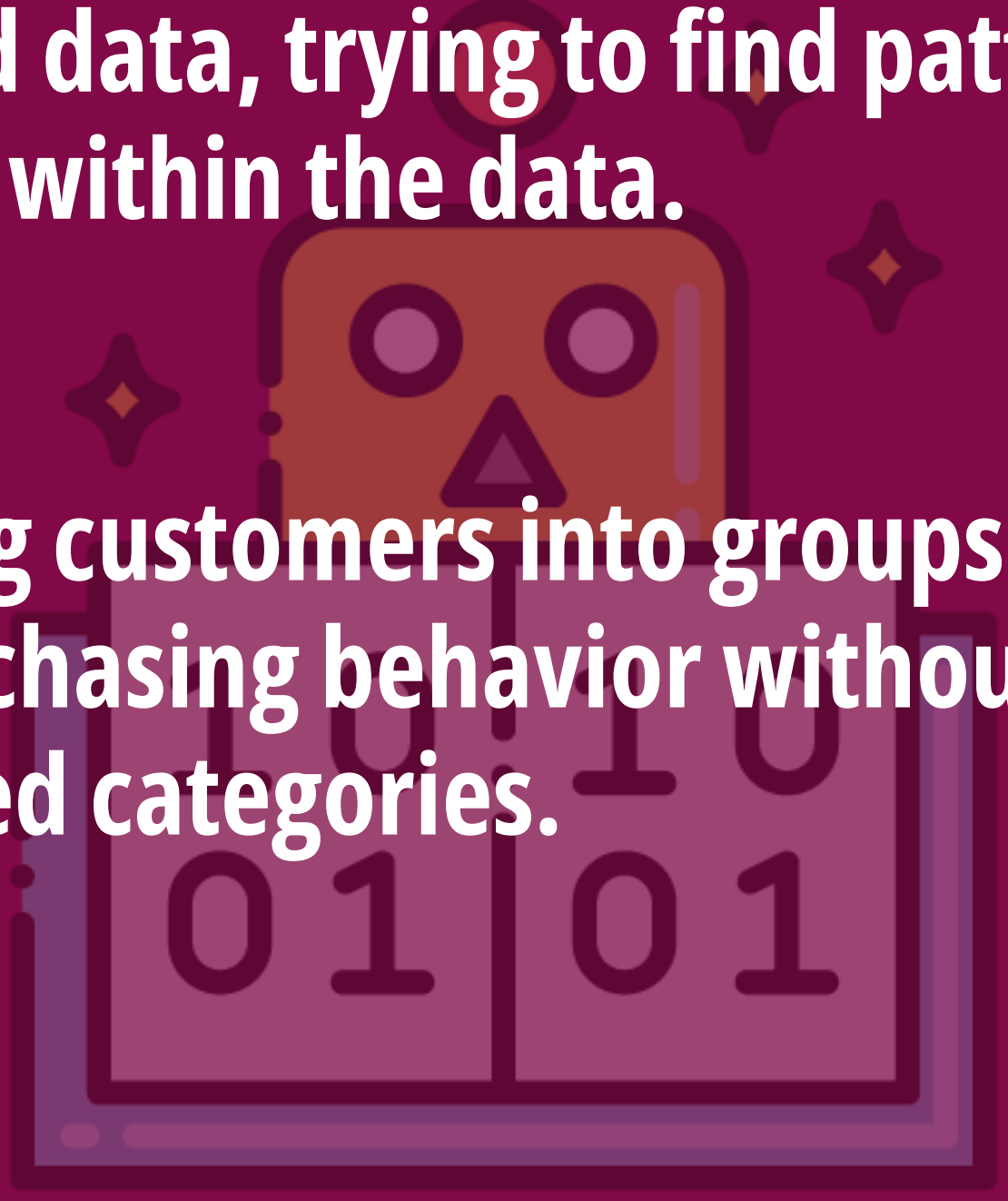
Training a model to recognize handwritten digits using a dataset of images labeled with the correct digit.

Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

9. Unsupervised Learning

A type of ML where the algorithm works on unlabeled data, trying to find patterns or structure within the data.



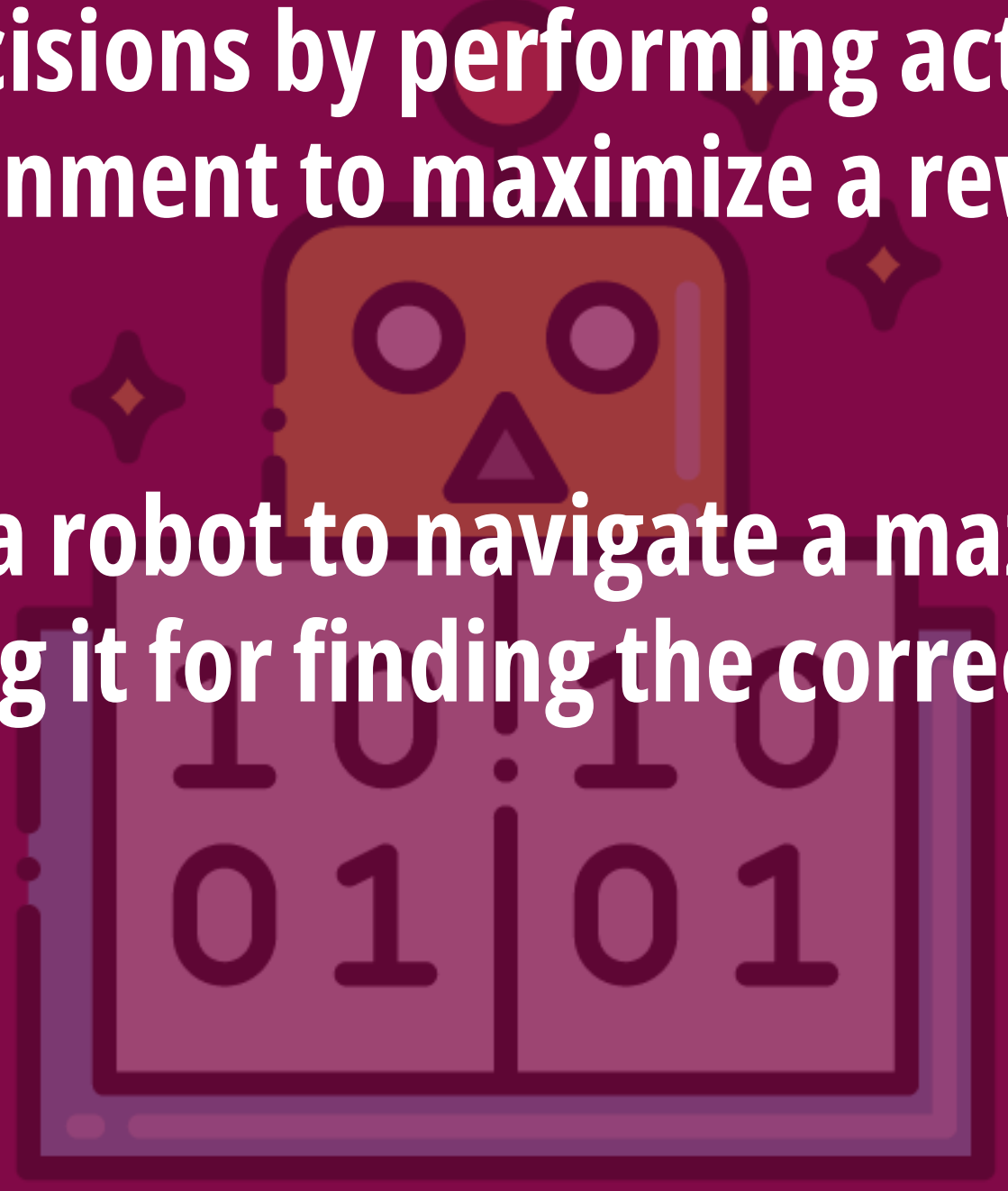
Clustering customers into groups based on their purchasing behavior without predefined categories.

Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

10. Reinforcement Learning

A type of ML where an agent learns to make decisions by performing actions in an environment to maximize a reward.



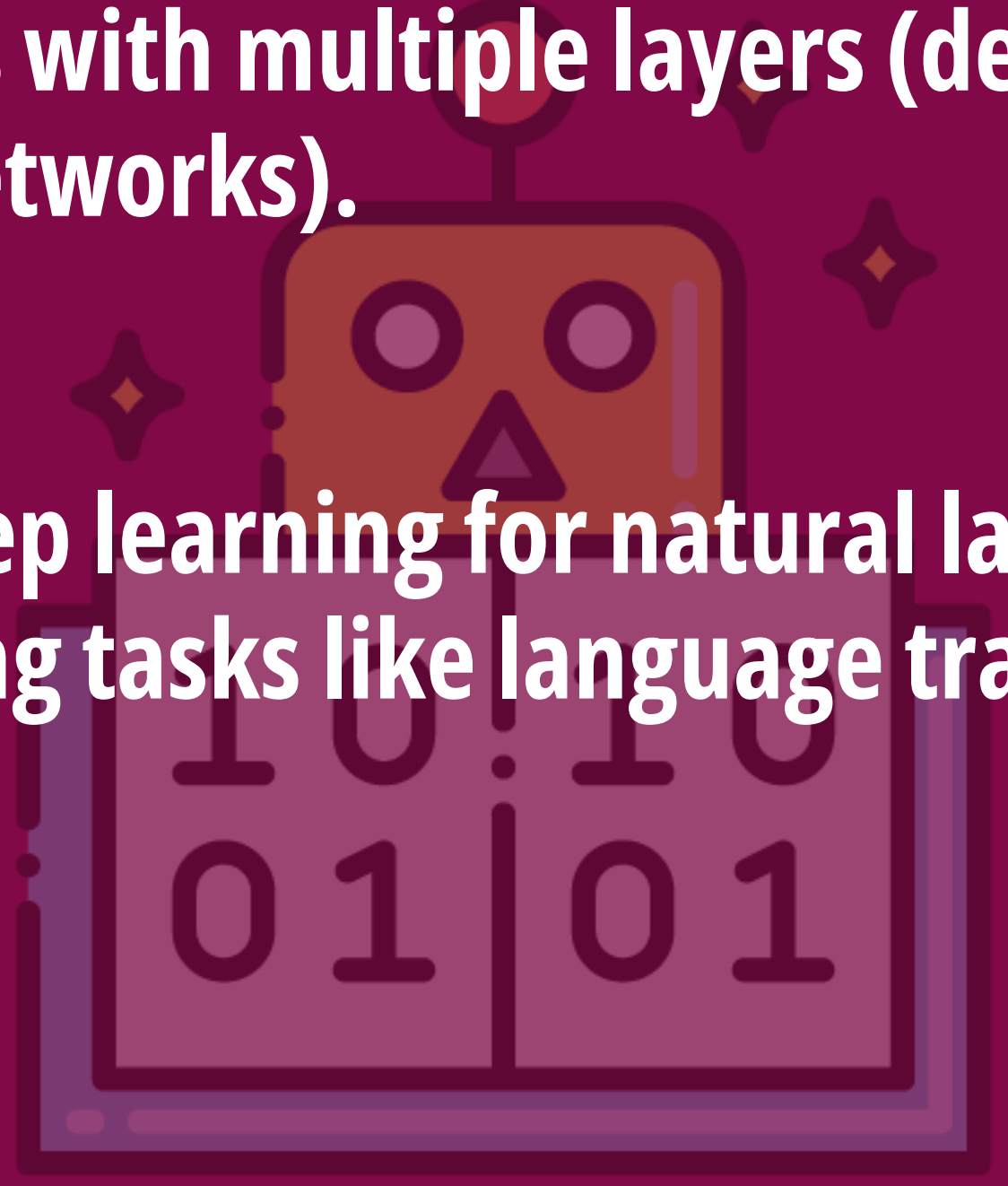
Training a robot to navigate a maze by rewarding it for finding the correct path.

Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

11. Deep Learning

A subset of ML based on artificial neural networks with multiple layers (deep neural networks).



Using deep learning for natural language processing tasks like language translation.

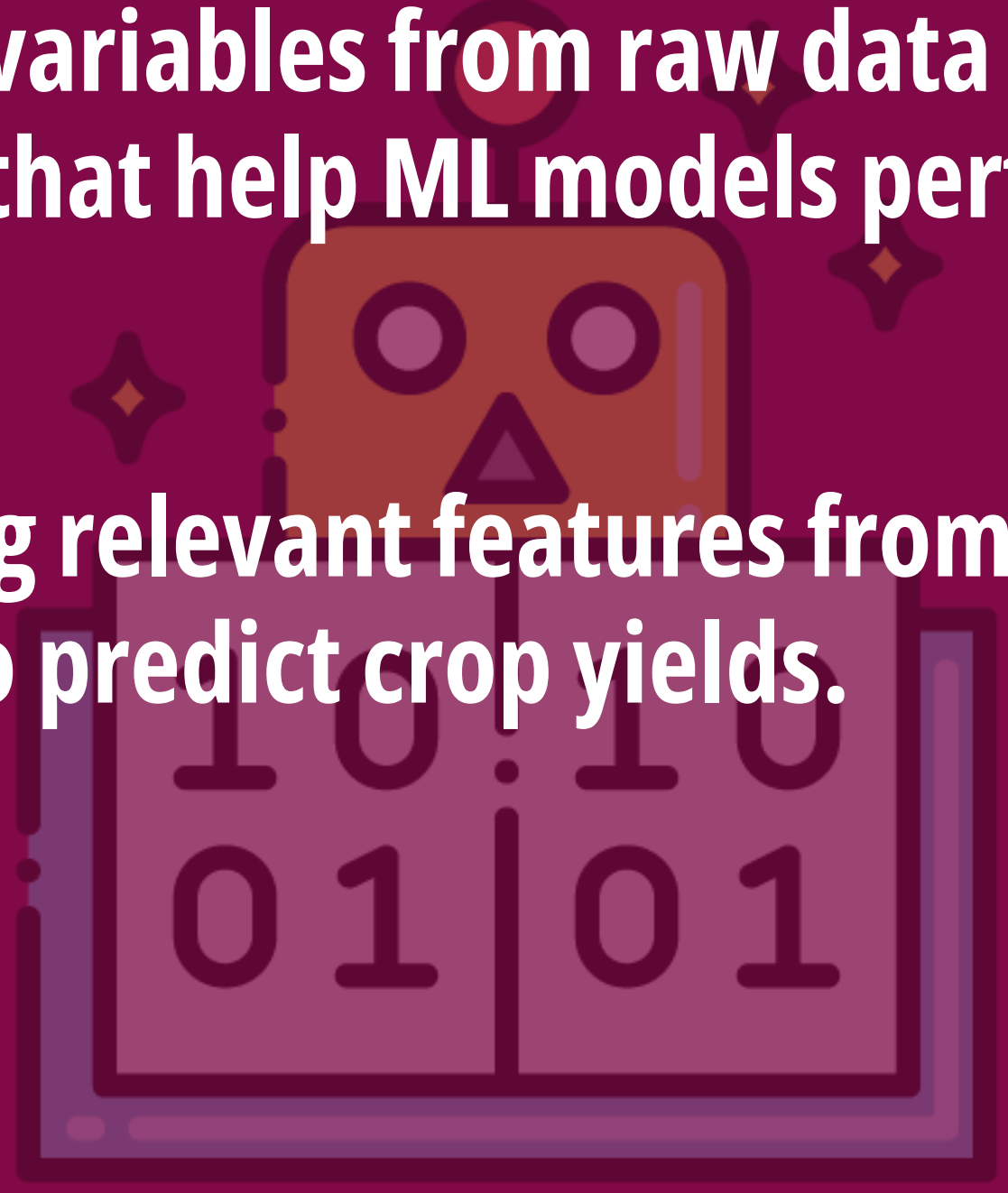
Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

12. Feature Engineering

The process of selecting and transforming relevant variables from raw data to create features that help ML models perform better.

Extracting relevant features from satellite images to predict crop yields.



Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

13. Model Training

The process of teaching a ML model to make predictions or decisions based on input data.

Training a sentiment analysis model on a large dataset of product reviews.

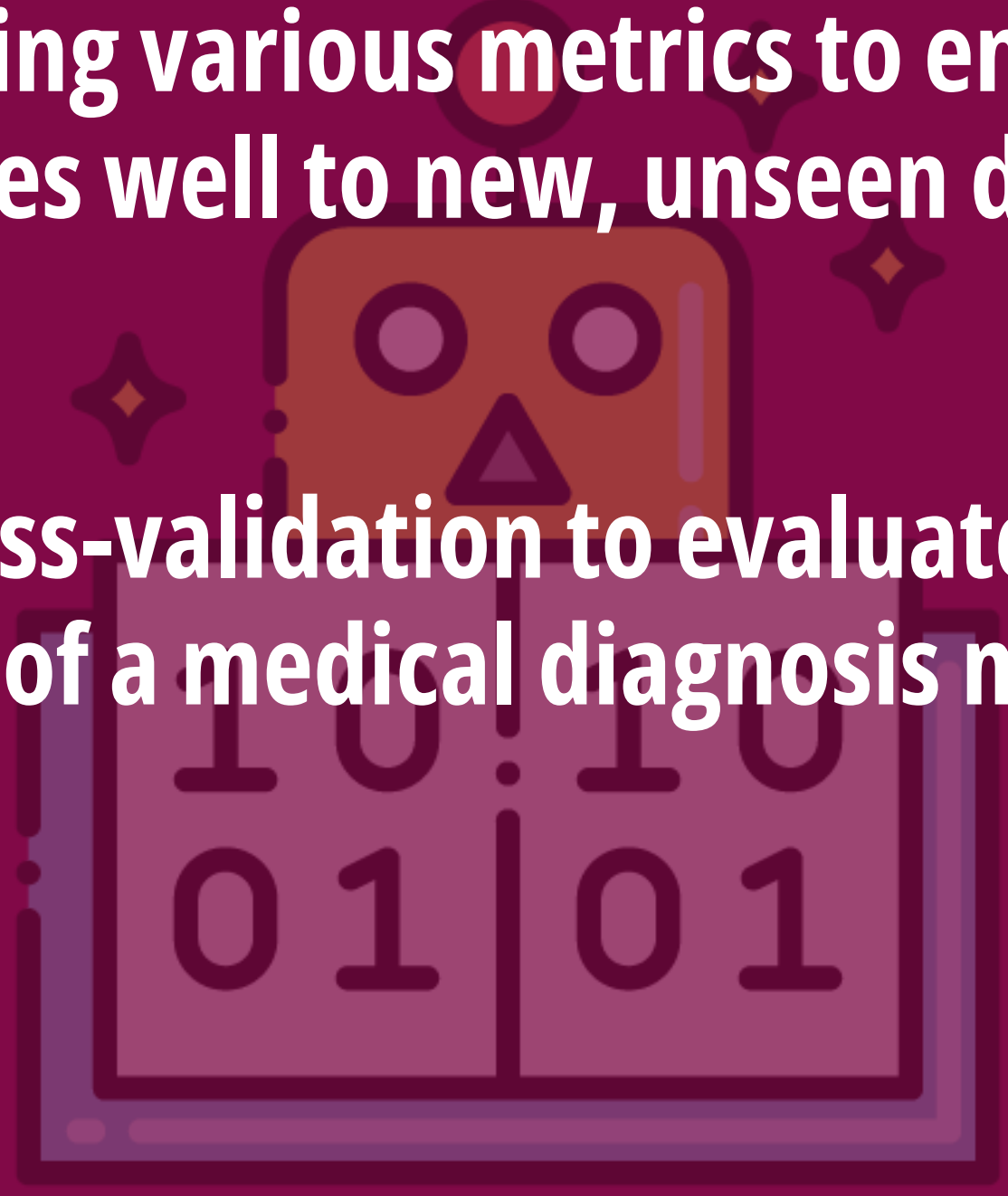
Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

14. Model Evaluation

Assessing the performance of a trained ML model using various metrics to ensure it generalizes well to new, unseen data.

Using cross-validation to evaluate the accuracy of a medical diagnosis model.



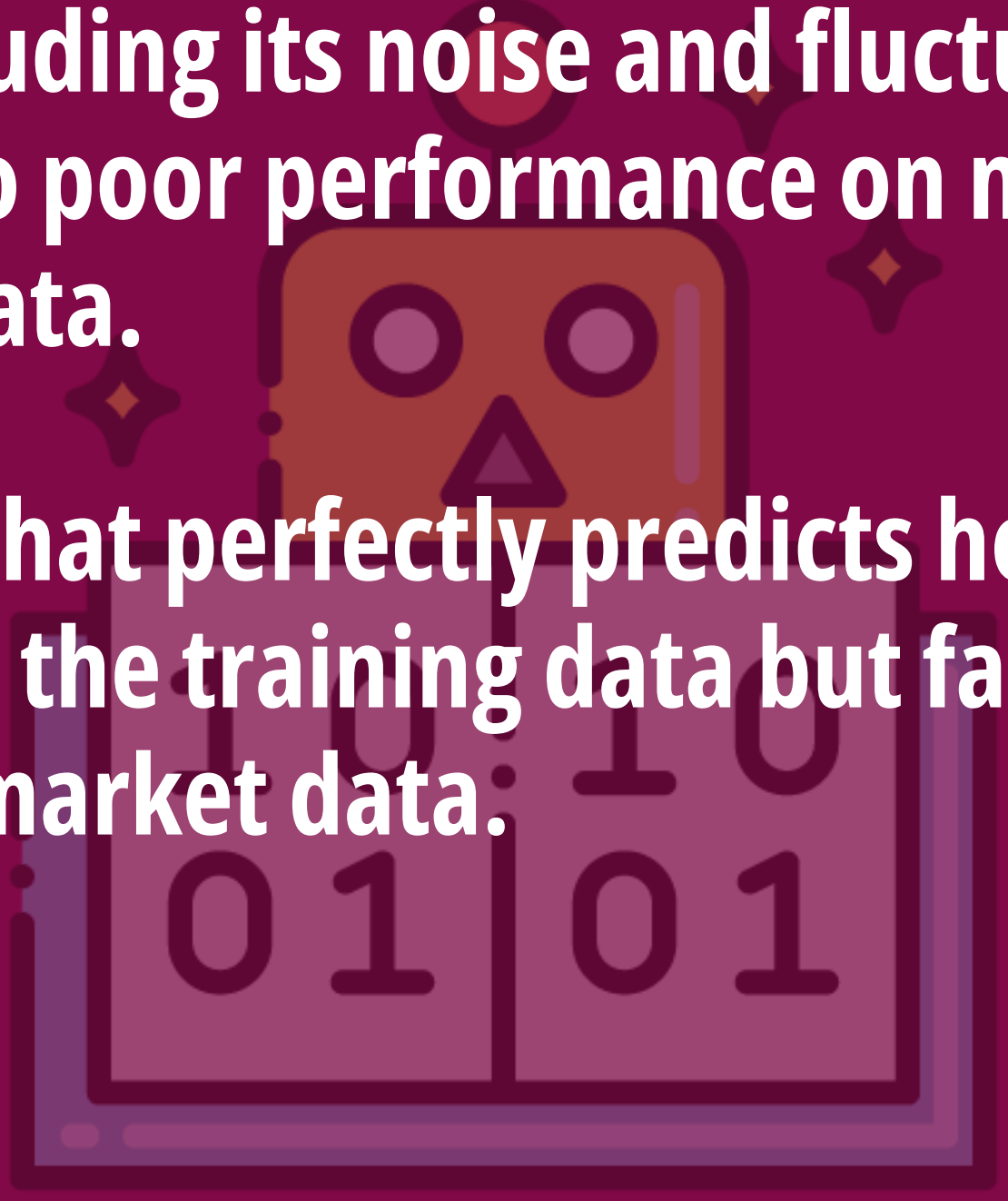
Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

15. Overfitting

When a model learns the training data too well, including its noise and fluctuations, leading to poor performance on new, unseen data.

A model that perfectly predicts house prices for the training data but fails on new housing market data.



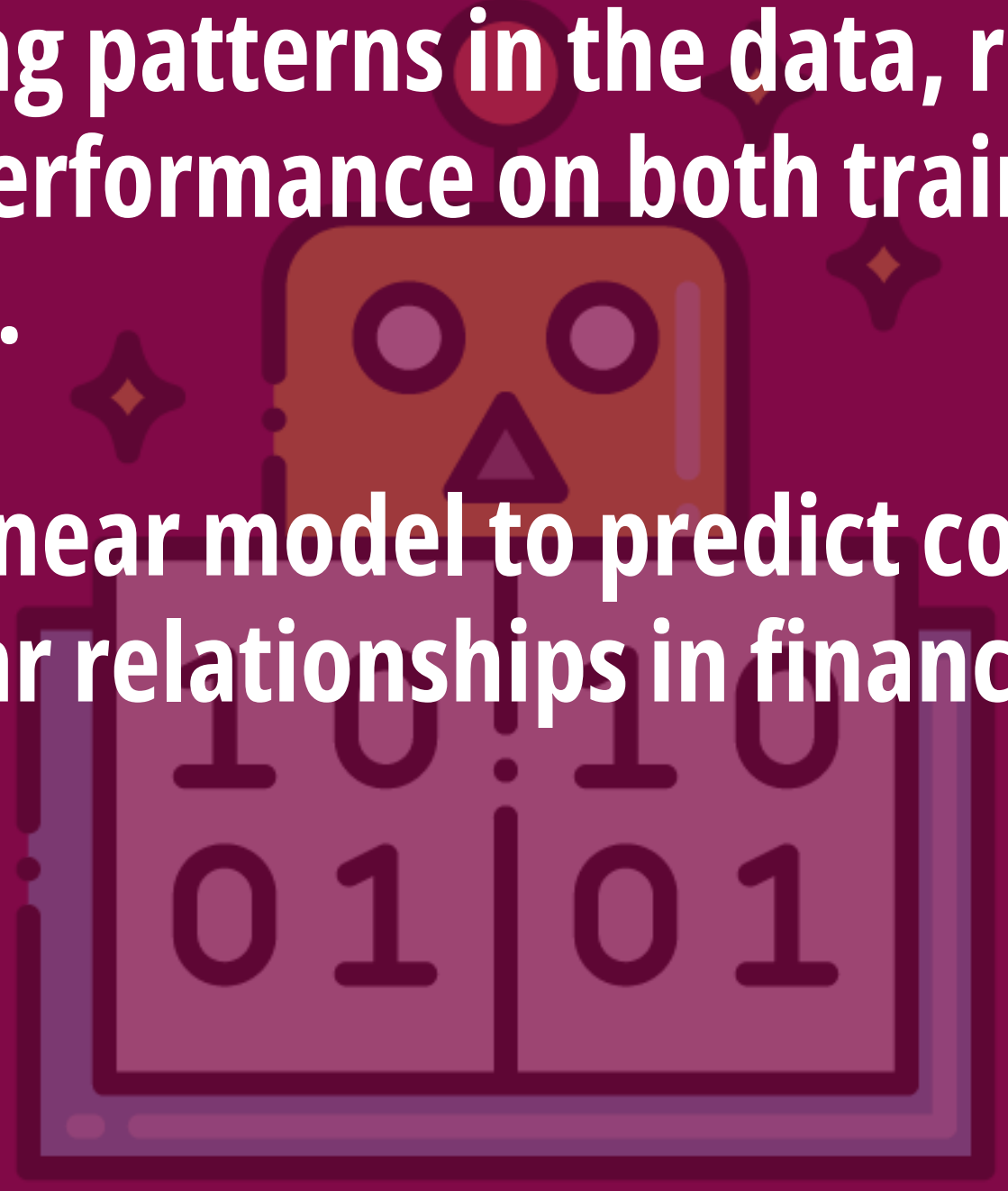
Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

16. Underfitting

When a model is too simple to capture the underlying patterns in the data, resulting in poor performance on both training and new data.

Using a linear model to predict complex, non-linear relationships in financial data.



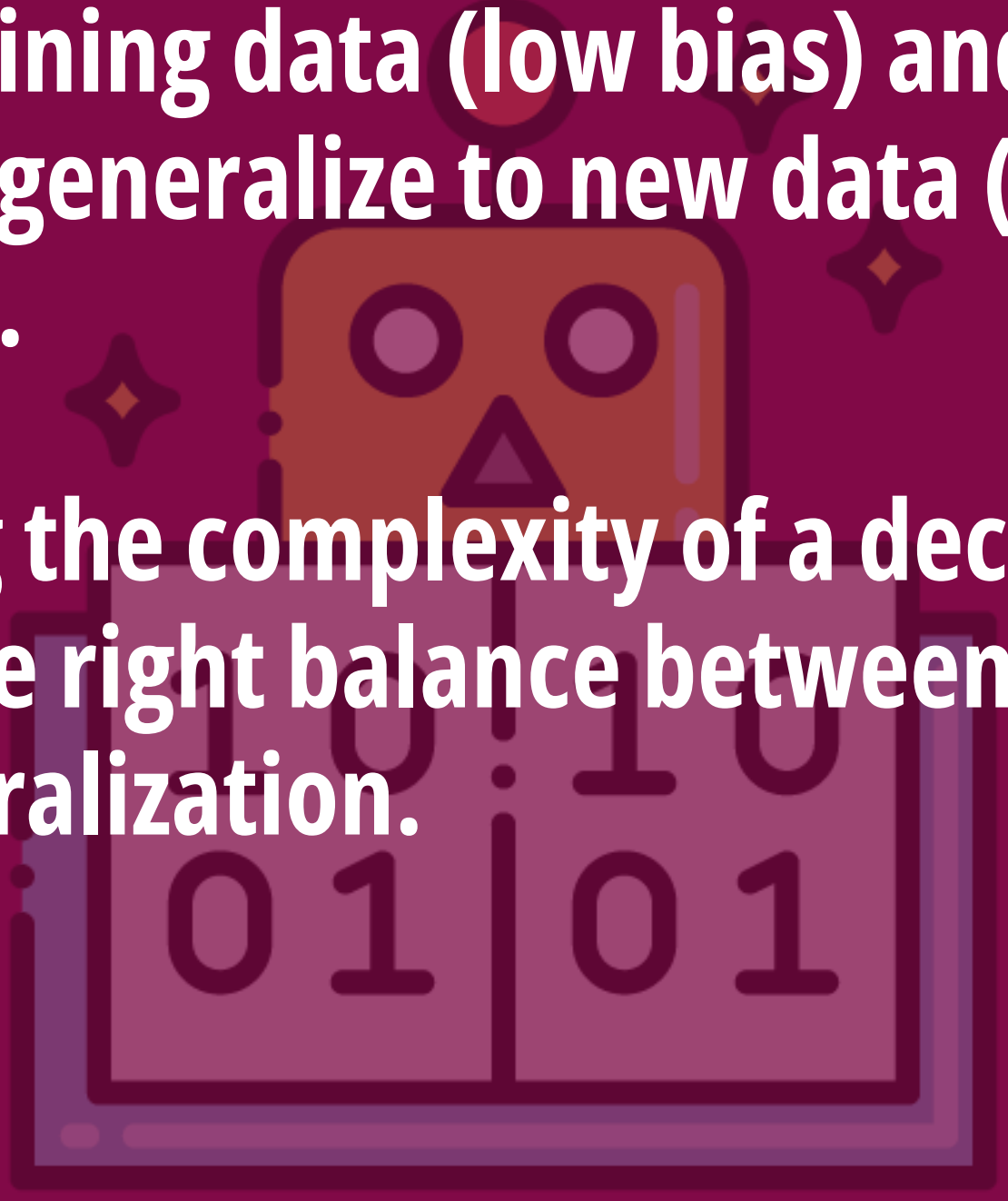
Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

17. Bias-Variance Tradeoff

The balance between a model's ability to fit the training data (low bias) and its ability to generalize to new data (low variance).

Adjusting the complexity of a decision tree to find the right balance between accuracy and generalization.



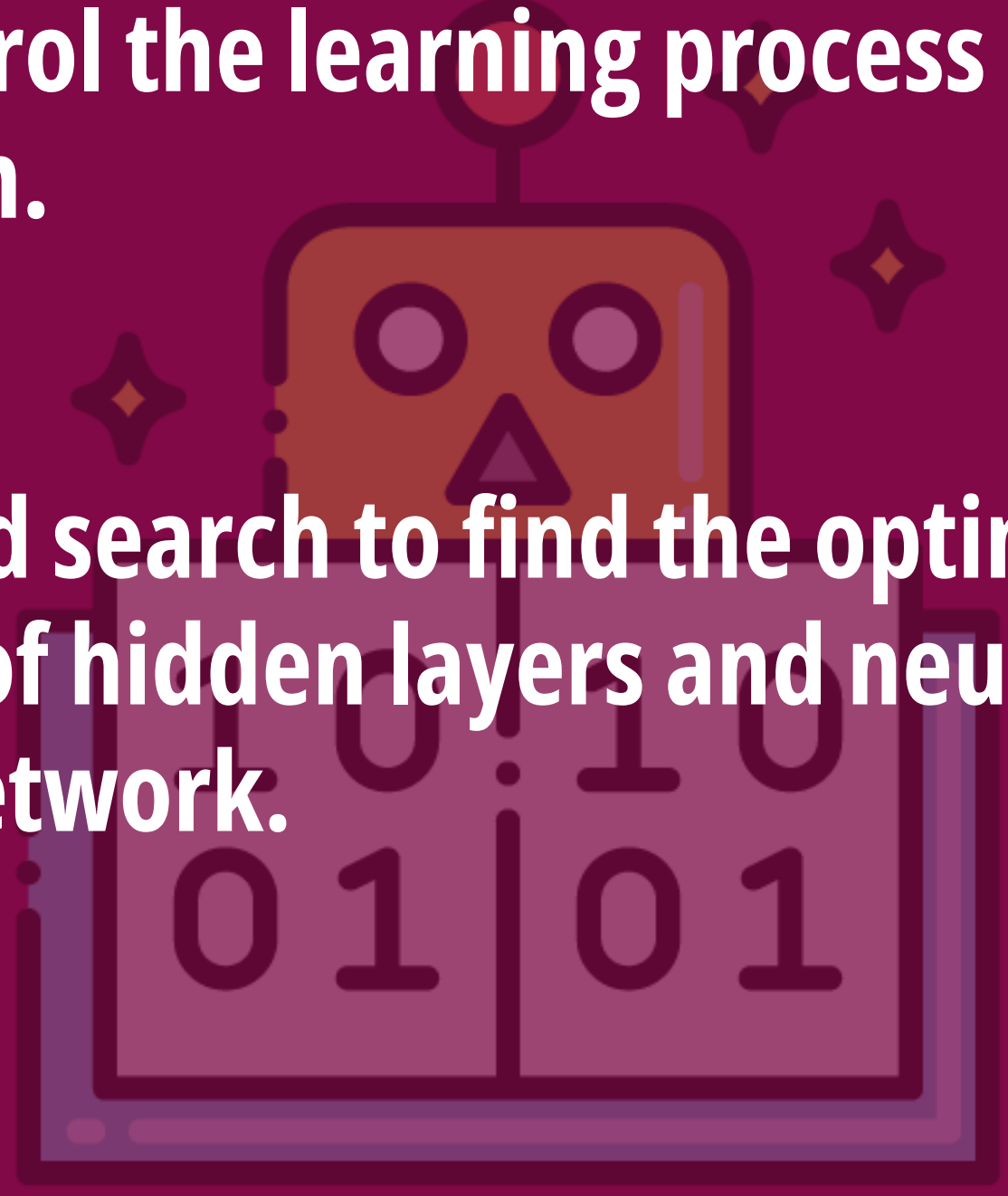
Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

18. Hyperparameter Tuning

The process of optimizing the parameters that control the learning process of an ML algorithm.

Using grid search to find the optimal number of hidden layers and neurons in a neural network.



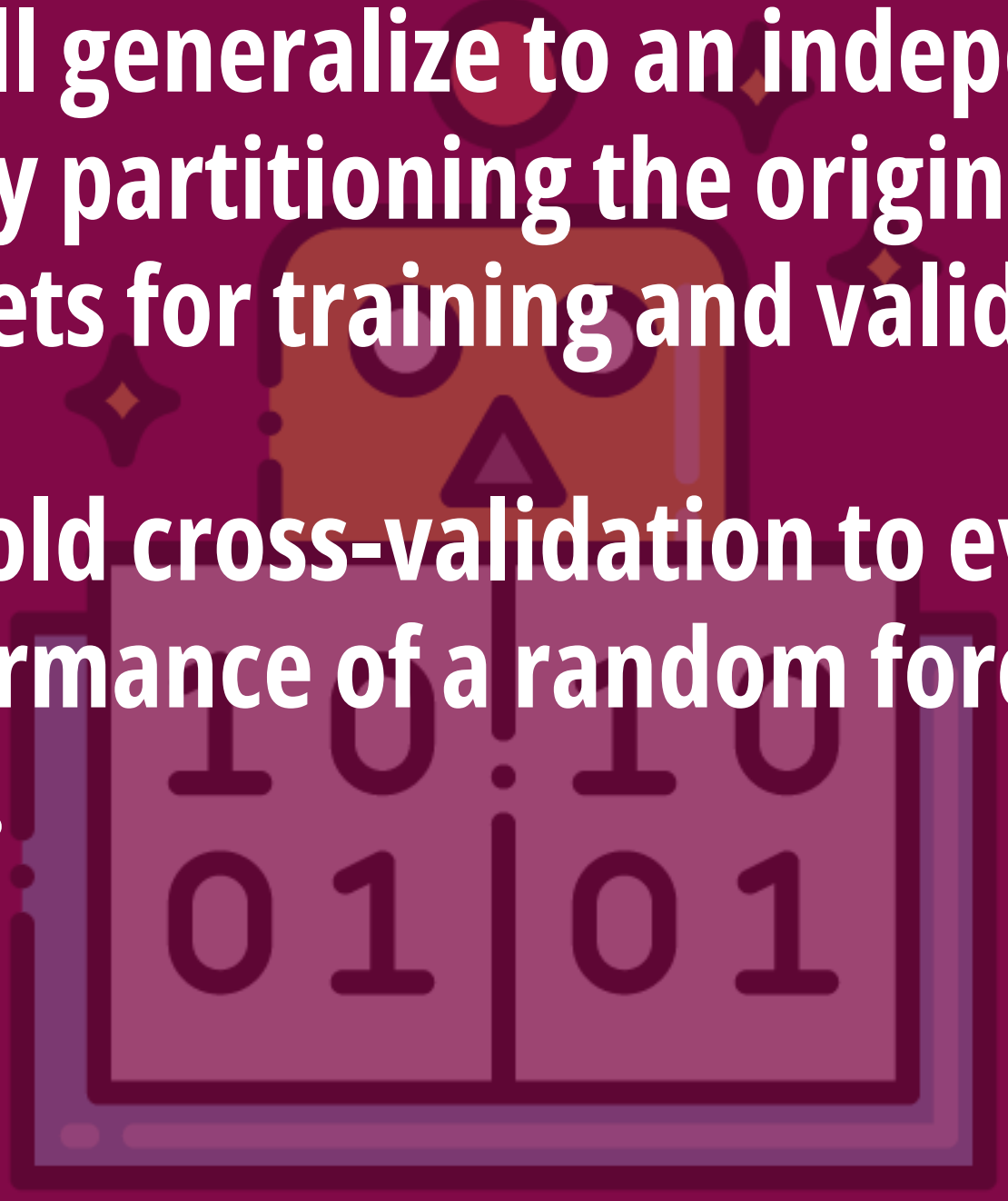
Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

19. Cross-validation

A technique for assessing how well a model will generalize to an independent dataset by partitioning the original sample into subsets for training and validation.

Using k-fold cross-validation to evaluate the performance of a random forest classifier.



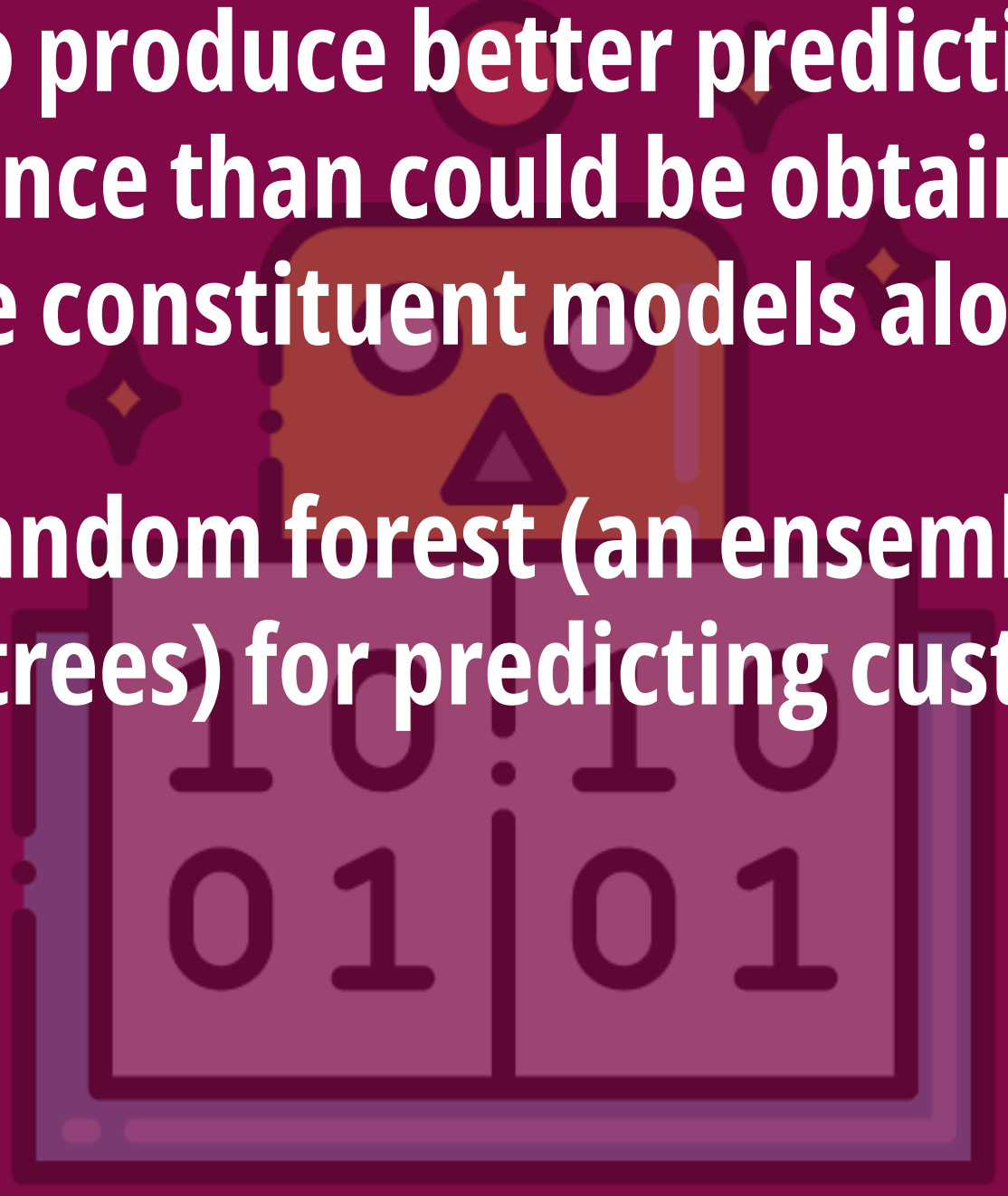
Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

20. Ensemble Methods

Techniques that combine multiple ML models to produce better predictive performance than could be obtained from any of the constituent models alone.

Using a random forest (an ensemble of decision trees) for predicting customer churn.

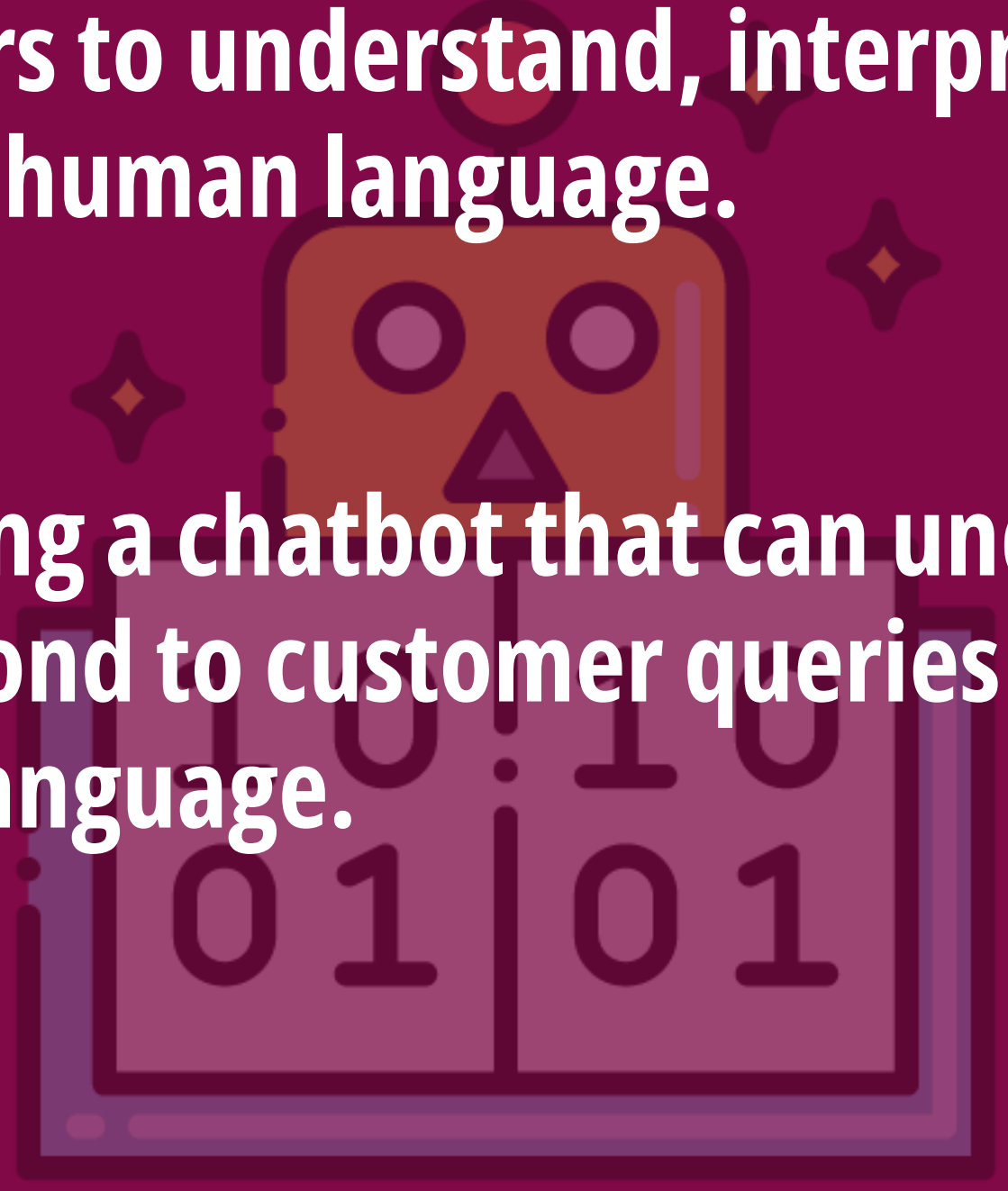


Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

21. Natural Language Processing (NLP)

A branch of ML focused on enabling computers to understand, interpret, and generate human language.



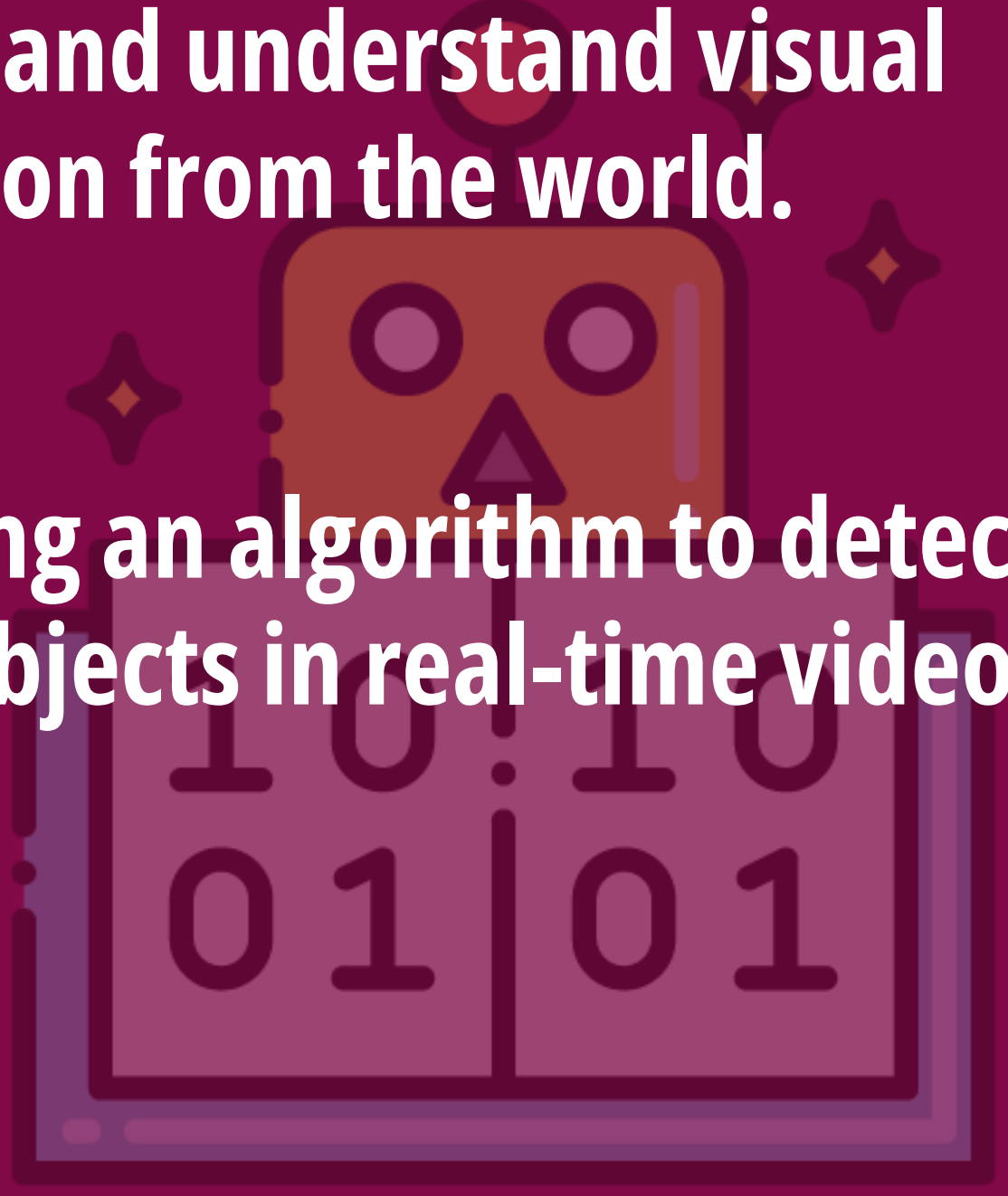
Developing a chatbot that can understand and respond to customer queries in natural language.

Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

22. Computer Vision

A field of ML that trains computers to interpret and understand visual information from the world.



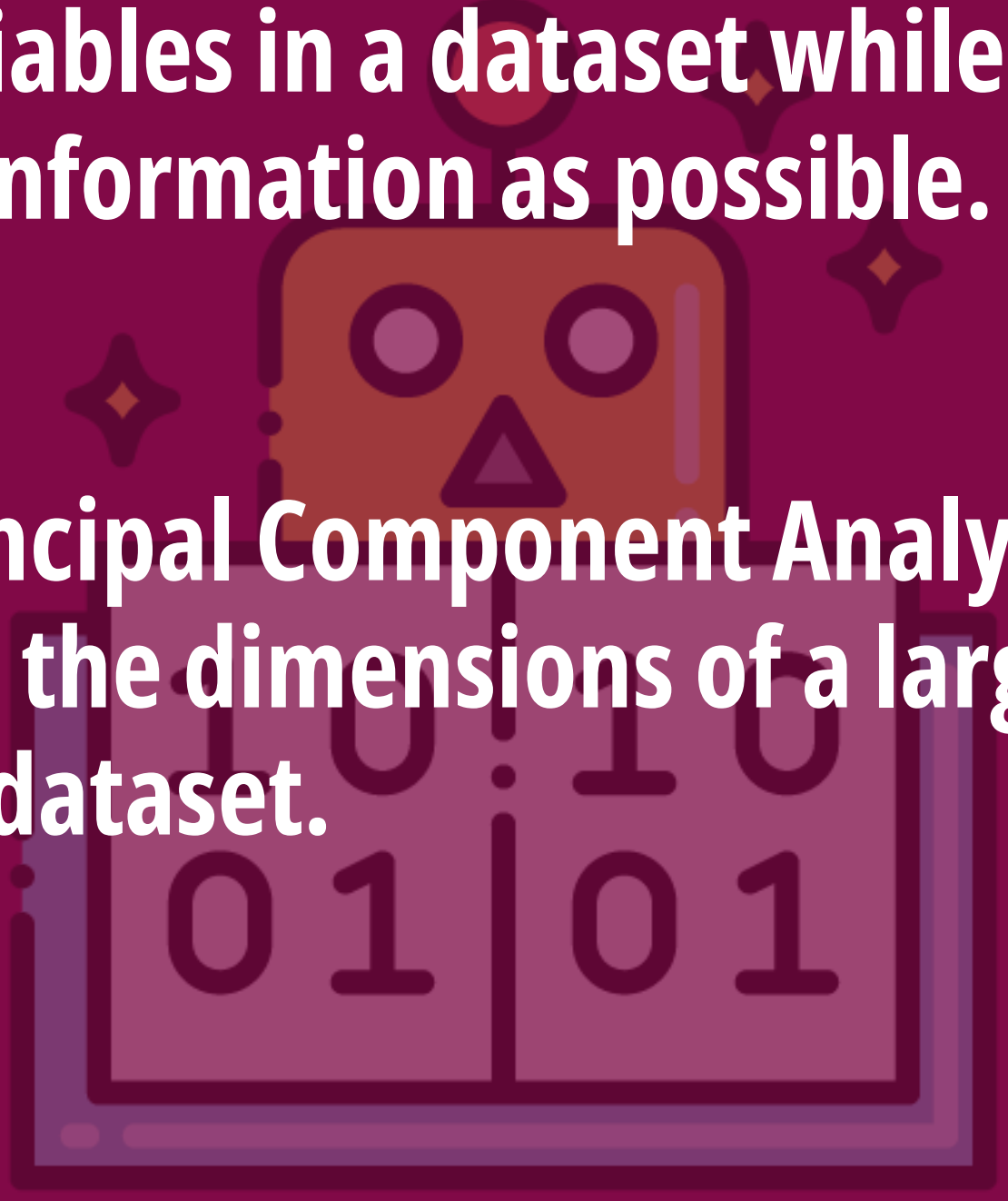
Developing an algorithm to detect and classify objects in real-time video streams.

Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

23. Dimensionality Reduction

Techniques for reducing the number of input variables in a dataset while retaining as much information as possible.



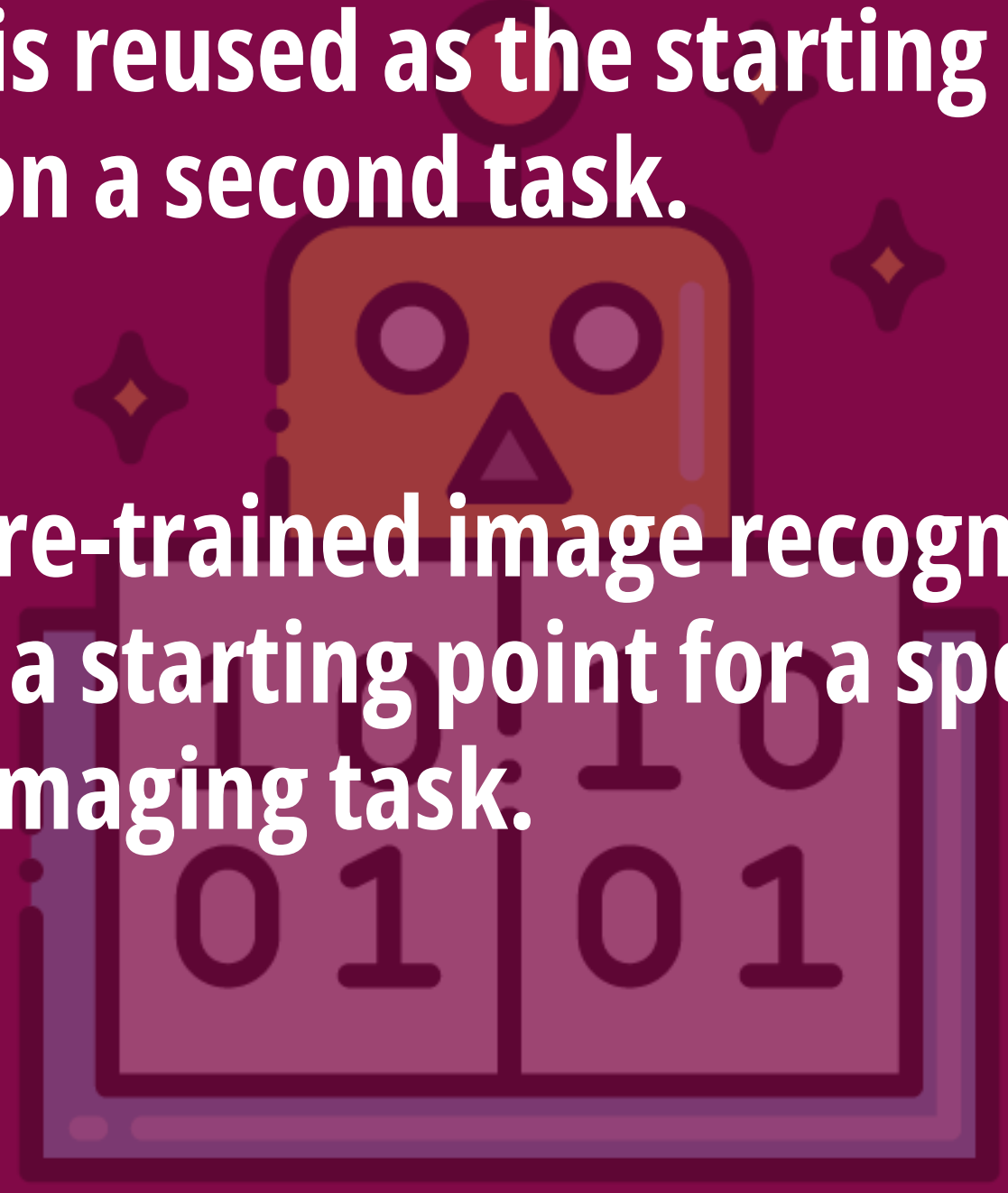
Using Principal Component Analysis (PCA) to reduce the dimensions of a large genomic dataset.

Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

24. Transfer Learning

A technique where a model developed for one task is reused as the starting point for a model on a second task.



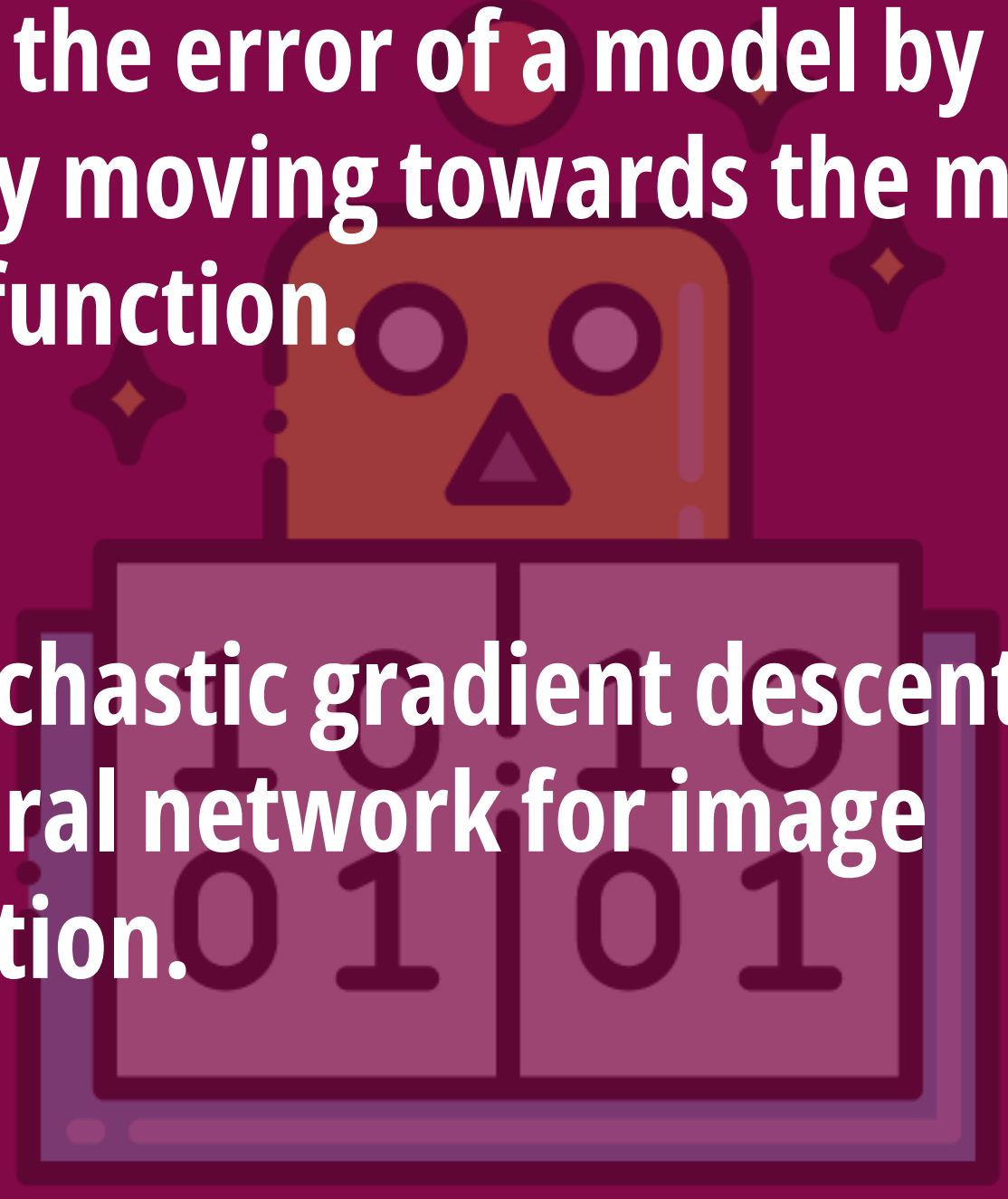
Using a pre-trained image recognition model as a starting point for a specific medical imaging task.

Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

25. Gradient Descent

An optimization algorithm used to minimize the error of a model by iteratively moving towards the minimum of a cost function.



Using stochastic gradient descent to train a large neural network for image classification.

Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

26. Regularization

Techniques used to prevent overfitting by adding a penalty term to the loss function.

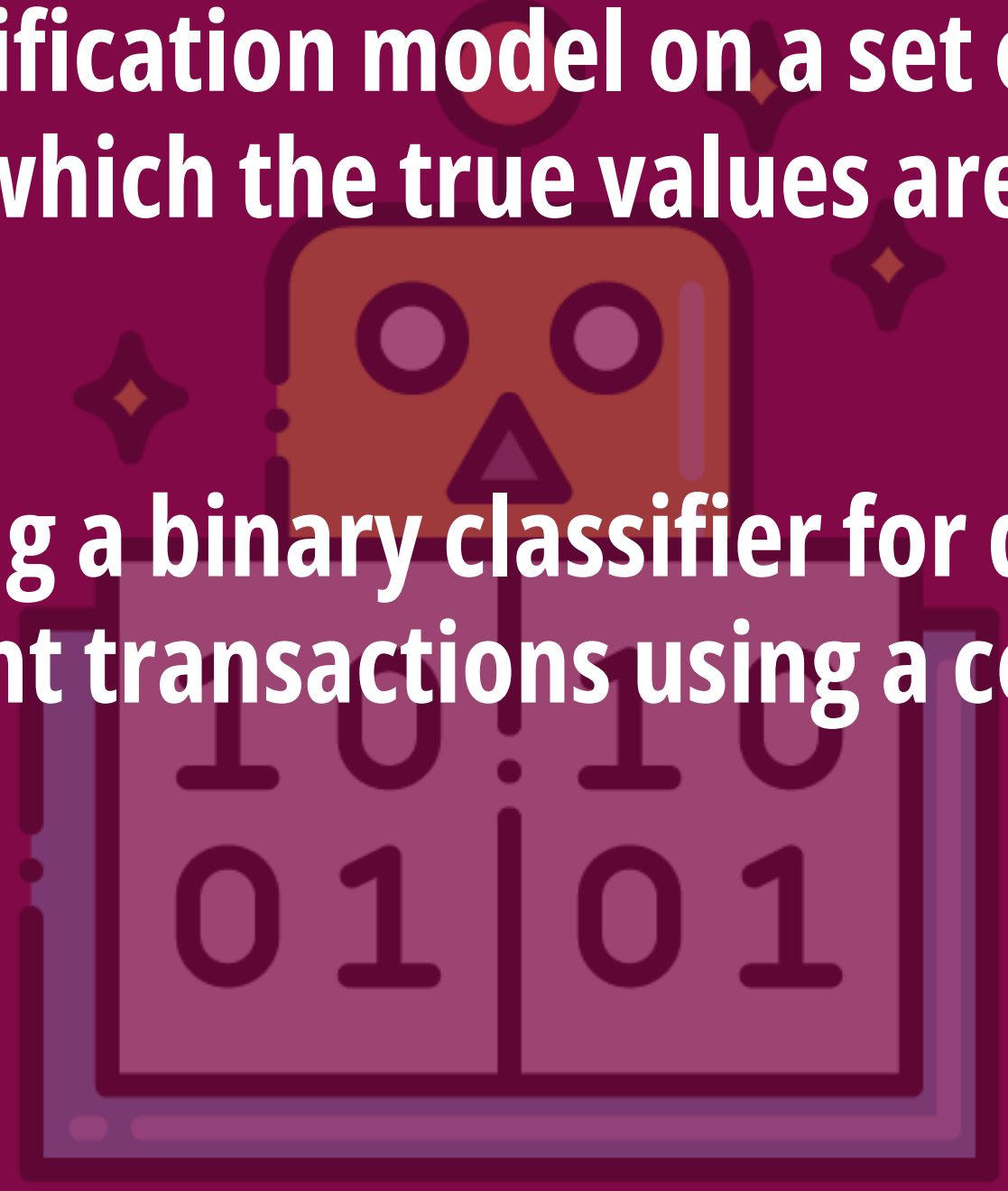
Applying L1 regularization (Lasso) to a linear regression model to perform feature selection.

Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

27. Confusion Matrix

A table used to describe the performance of a classification model on a set of test data for which the true values are known.



Evaluating a binary classifier for detecting fraudulent transactions using a confusion matrix.

Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

28. ROC Curve

A graphical plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied.

Comparing the performance of different models for diagnosing a disease using ROC curves.

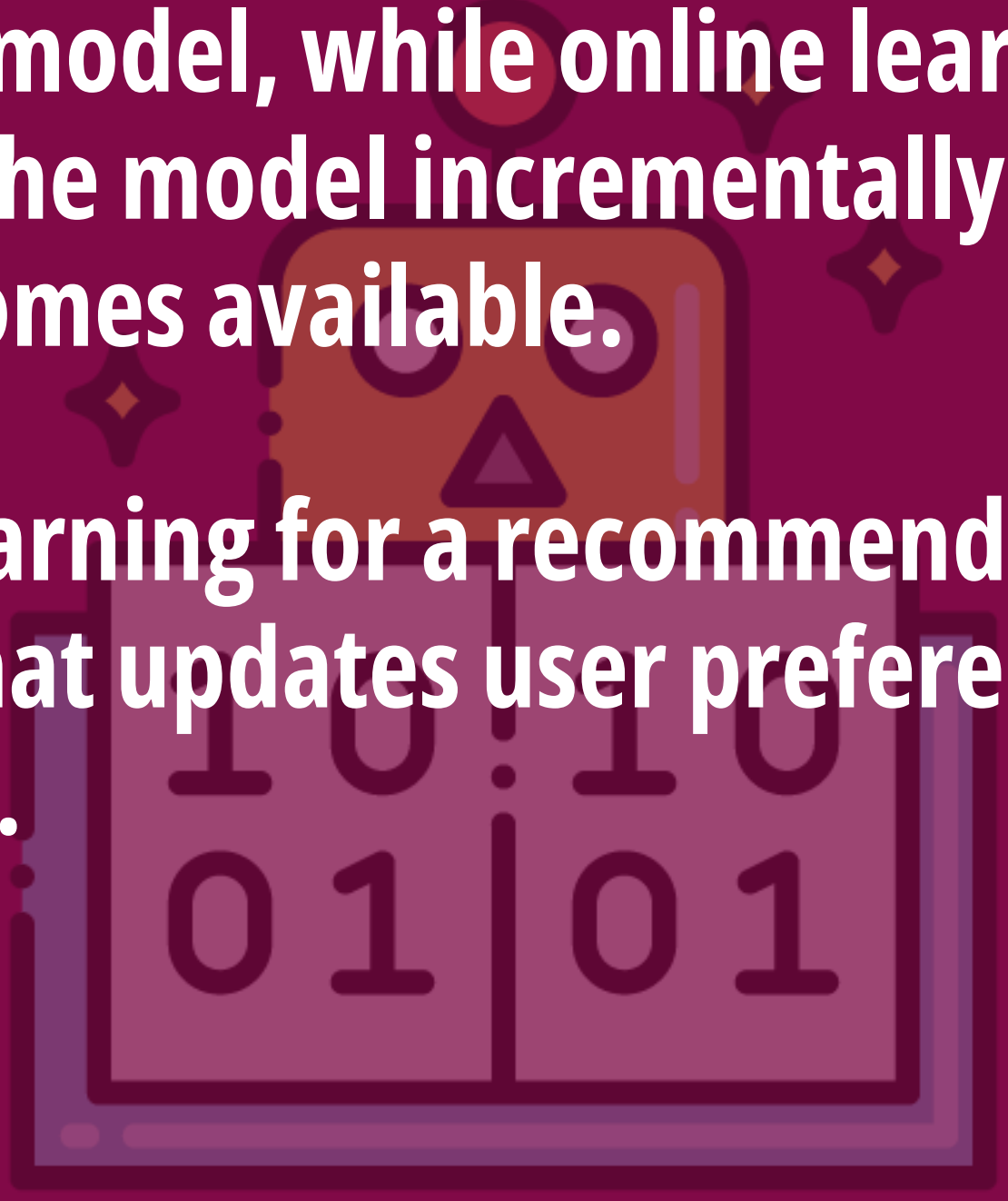
Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

29. Batch vs. Online Learning

Batch learning uses all available data to train the model, while online learning updates the model incrementally as new data becomes available.

Online learning for a recommendation system that updates user preferences in real-time.

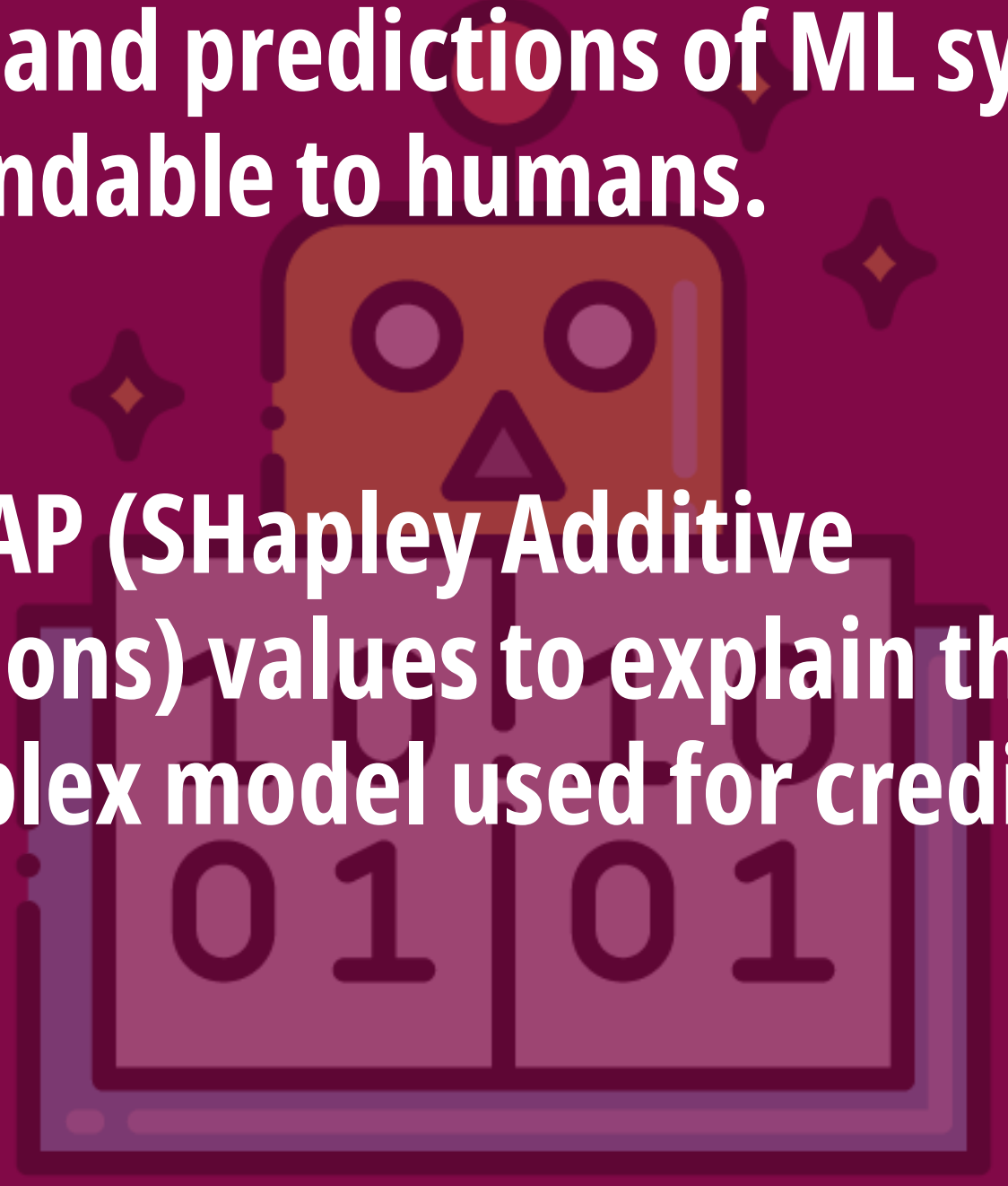


Shwetank Singh
GritSetGrow - GSGLearn.com

Machine Learning 101: Basics

30. Explainable AI

Techniques and methods that make the behavior and predictions of ML systems understandable to humans.



Using SHAP (SHapley Additive exPlanations) values to explain the output of a complex model used for credit scoring.

Shwetank Singh
GritSetGrow - GSGLearn.com

Thank you