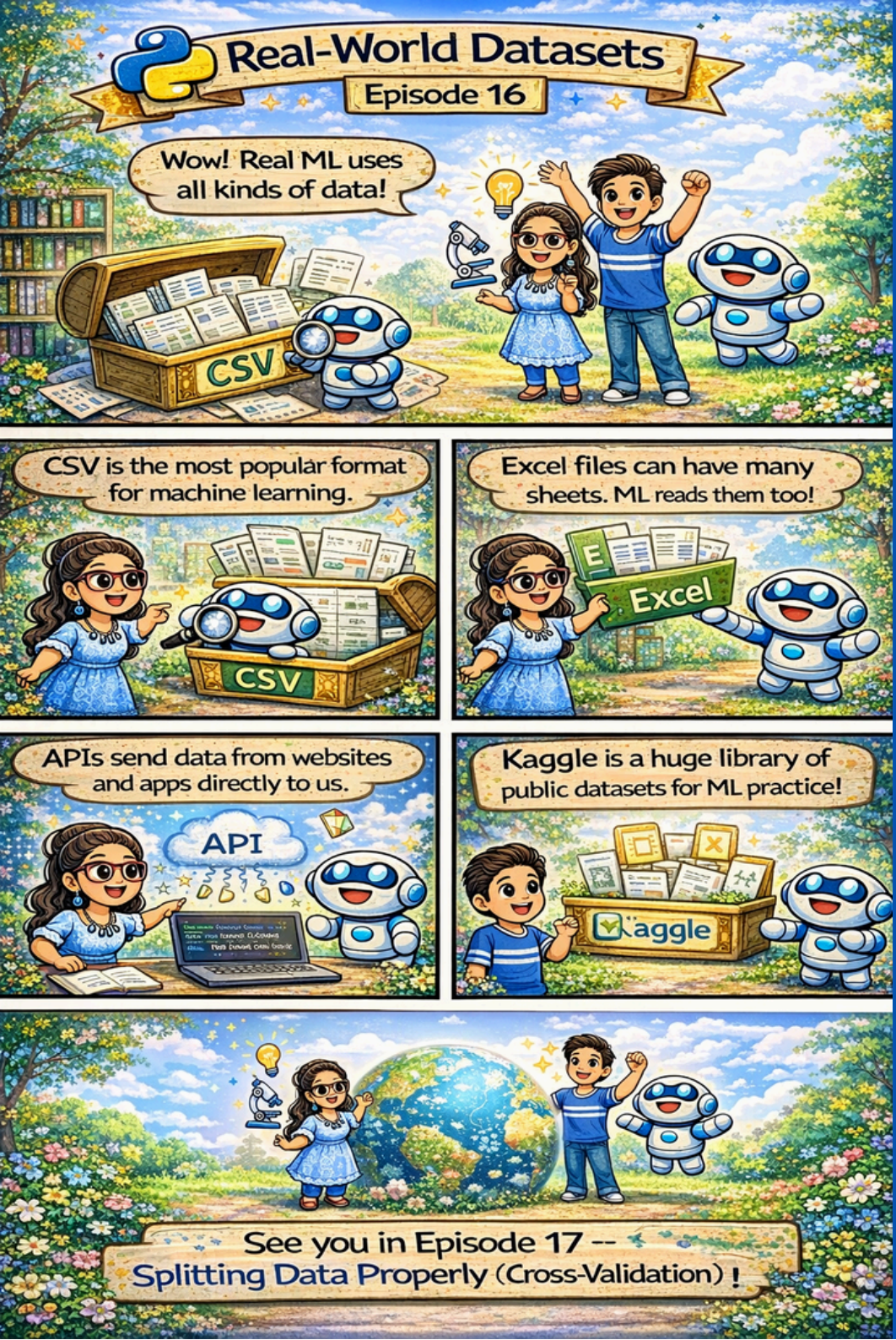
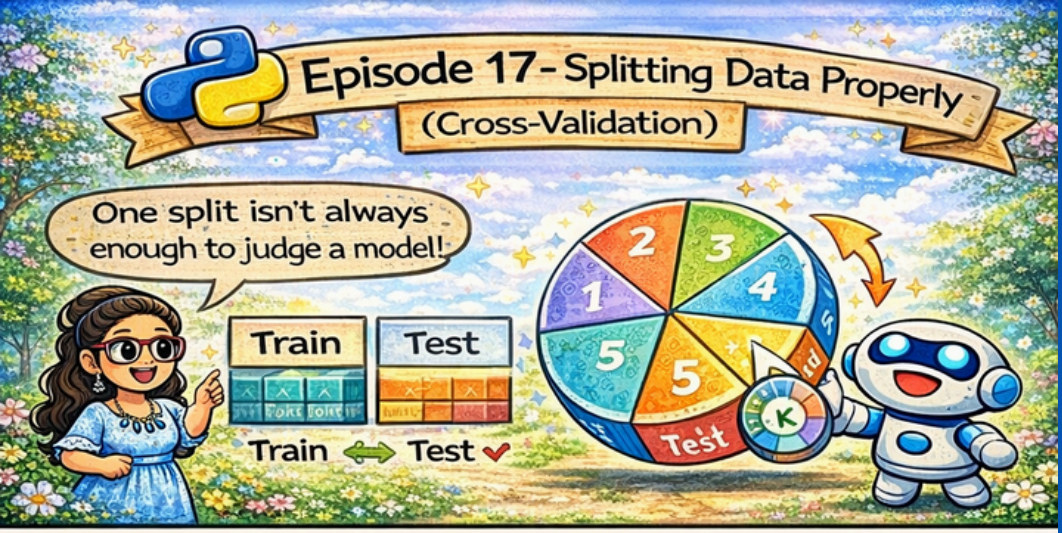
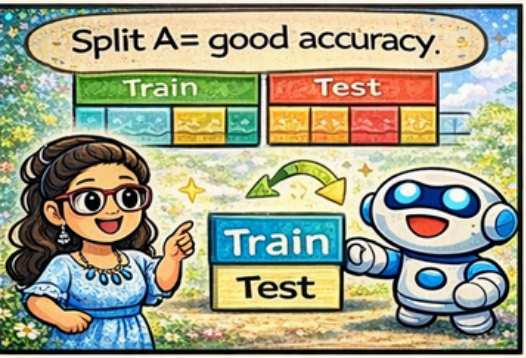
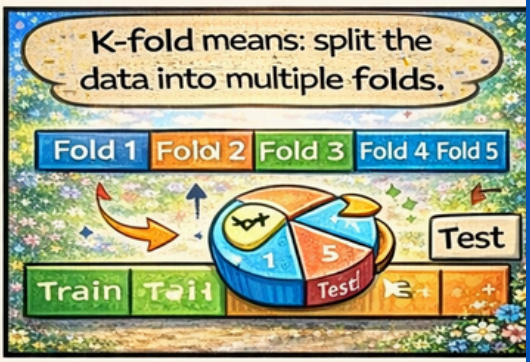

MACHINE LEARNING SERIES!
AI & ML Adventures for Kids!

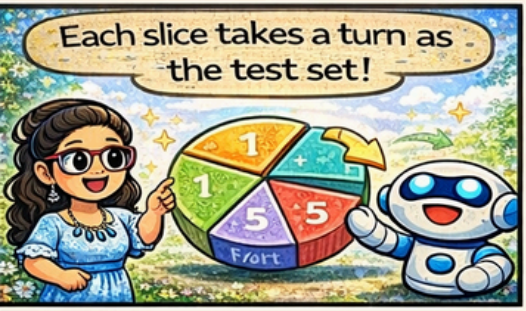

Real-World Datasets
Episode 16
Wow! Real ML uses all kinds of data!
CSV
CSV is the most popular format for machine learning.
CSV
Excel files can have many sheets. ML reads them too!
Excel
APIs send data from websites and apps directly to us.
API
Kaggle is a huge library of public datasets for ML practice!
Kaggle
See you in Episode 17 --
Splitting Data Properly (Cross-Validation) !

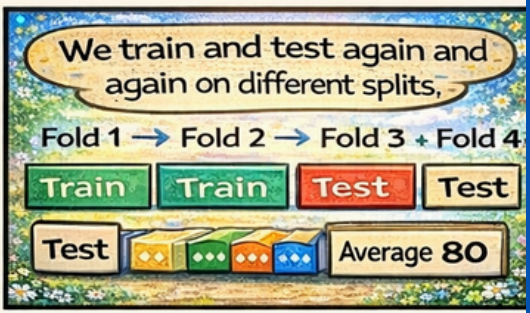
Episode 17 - Splitting Data Properly (Cross-Validation)



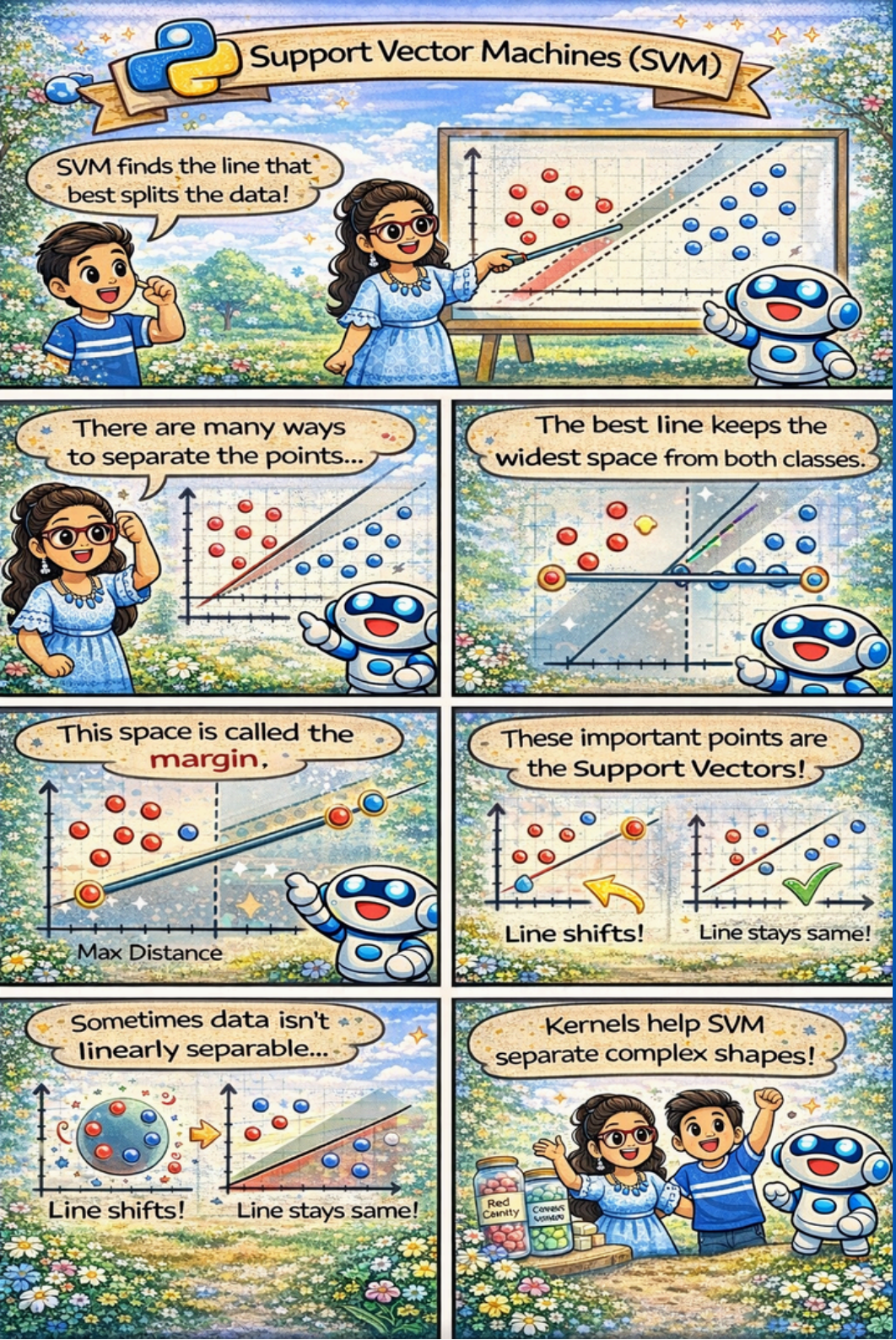


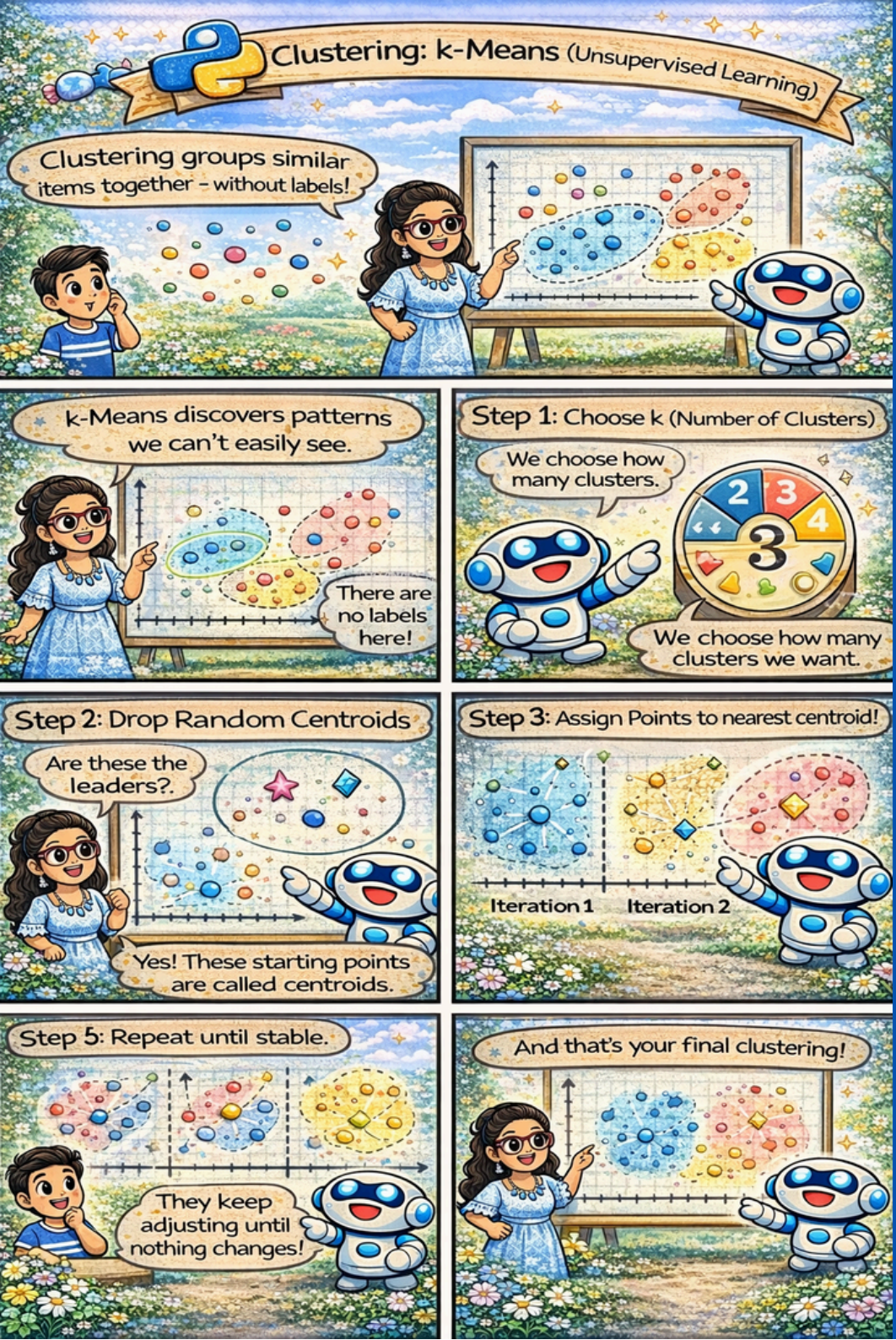


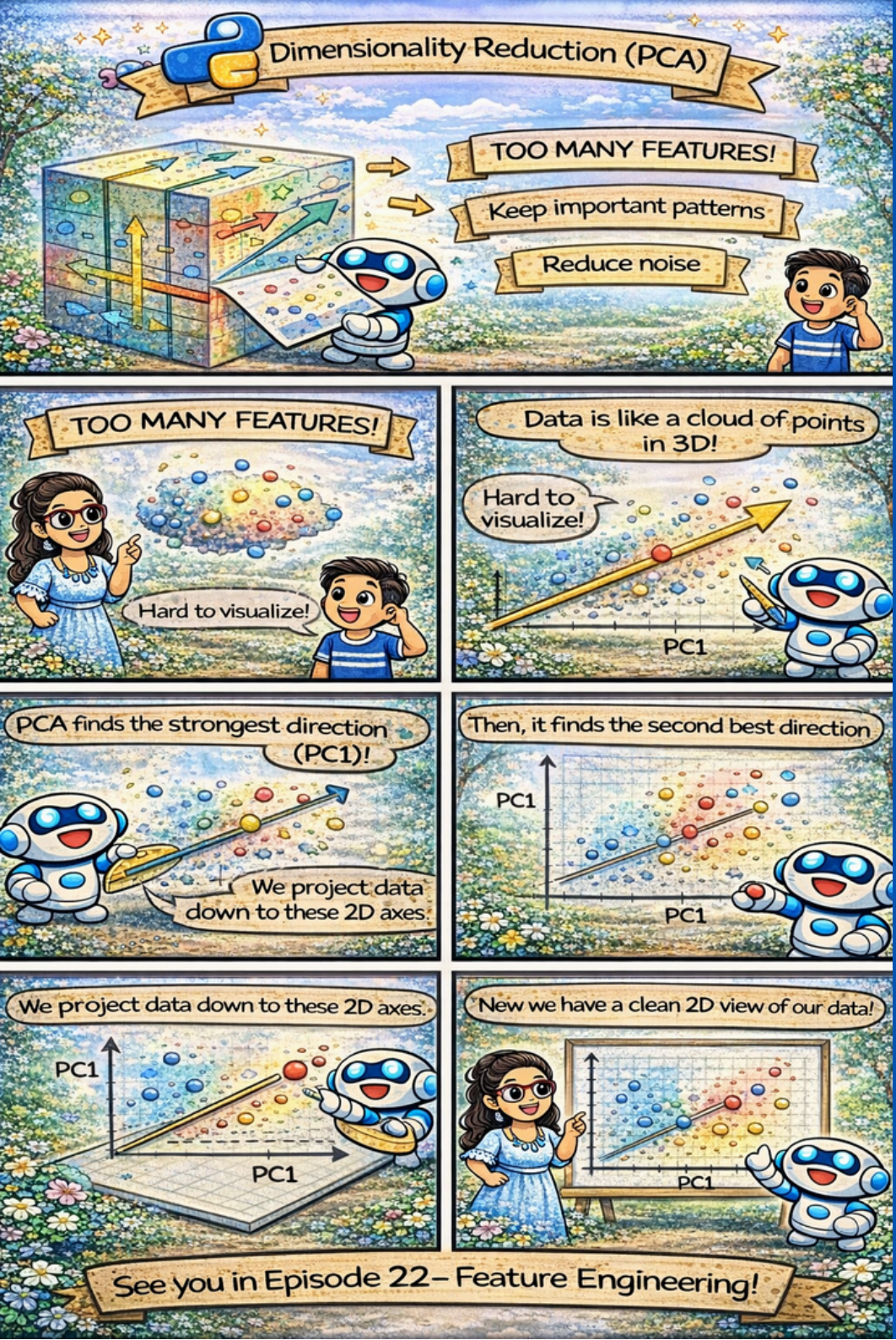


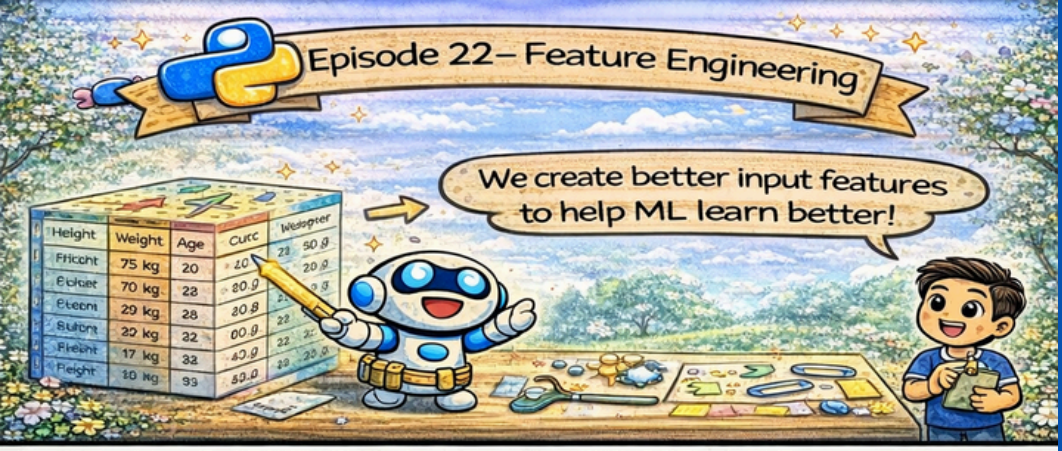


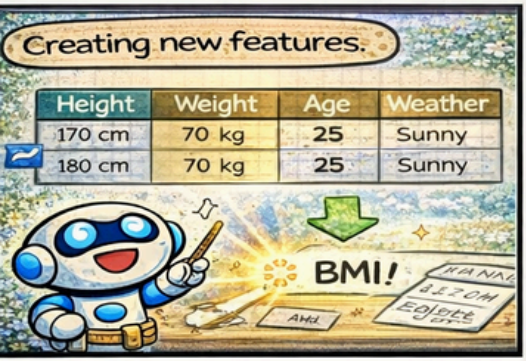

Episode 18 – Naive Bayes
Probability made simple!
Naive Bayes predicts using probabilities!
Red Candy
Blue Candy
Green Candy
If most candies are red...
Red
Blue
Green
Count
Naive Bayes starts by counting how many times each class appears.
Color
Flavor
Shape
Blue
Green
0.6
Each feature helps update the probability.
Color
Flavor
Shape
All probabilities combine to pick the best class!
Red: 0.6
Blue: 0.3
Green: 0.1
Red
It's simple AND powerful!
Text Classification
Spam Detection
Simple, Fast Predictions
Green

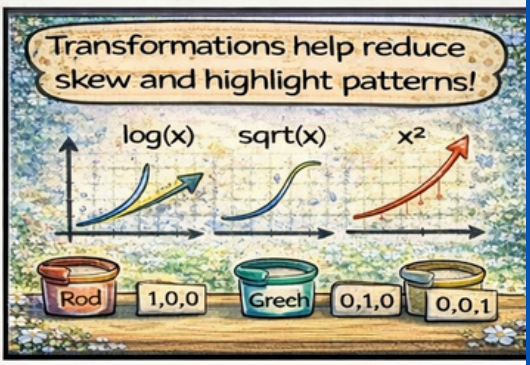

Support Vector Machines (SVM)
SVM finds the line that best splits the data!
There are many ways to separate the points...
The best line keeps the widest space from both classes.
This space is called the margin.
Max Distance
These important points are the Support Vectors!
Line shifts!
Line stays same!
Sometimes data isn't linearly separable...
Line shifts!
Line stays same!
Kernels help SVM separate complex shapes!

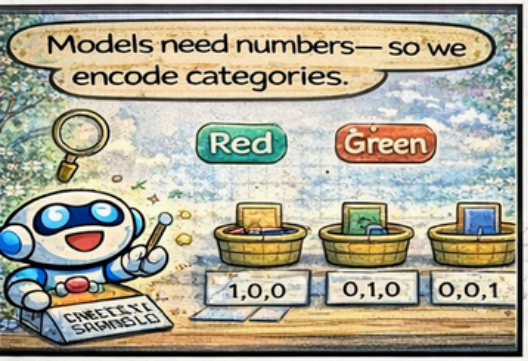

Clustering: k-Means (Unsupervised Learning)
Clustering groups similar items together - without labels!
k-Means discovers patterns we can't easily see.
There are no labels here!
Step 1: Choose k (Number of Clusters)
We choose how many clusters.
2
3
4
3
We choose how many clusters we want.
Step 2: Drop Random Centroids
Are these the leaders?.
Yes! These starting points are called centroids.
Step 3: Assign Points to nearest centroid!
Iteration 1
Iteration 2
Step 5: Repeat until stable.
They keep adjusting until nothing changes!
And that's your final clustering!


Dimensionality Reduction (PCA)
TOO MANY FEATURES!
Keep important patterns
Reduce noise
TOO MANY FEATURES!
Hard to visualize!
Data is like a cloud of points in 3D!
Hard to visualize!
PC1
PCA finds the strongest direction (PC1)!
We project data down to these 2D axes.
Then, it finds the second best direction
PC1
PC1
We project data down to these 2D axes.
PC1
PC1
New we have a clean 2D view of our data!
PC1
See you in Episode 22– Feature Engineering!

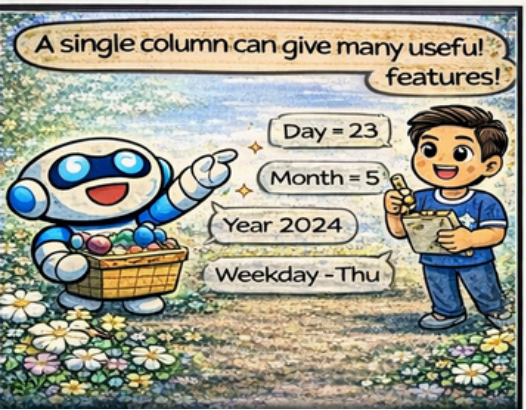

Episode 22– Feature Engineering
We create better input features to help ML learn better!

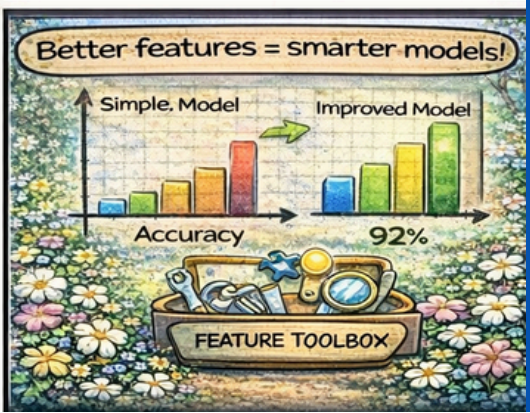

Creating new features.
Height
Weight
Age
Weather
170 cm
70 kg
25
Sunny
180 cm
70 kg
25
Sunny
BMI!


Transformations help reduce skew and highlight patterns!
log(x)
sqrt(x)
x²
Rod
1,0,0
Grech
0,1,0
0,0,1

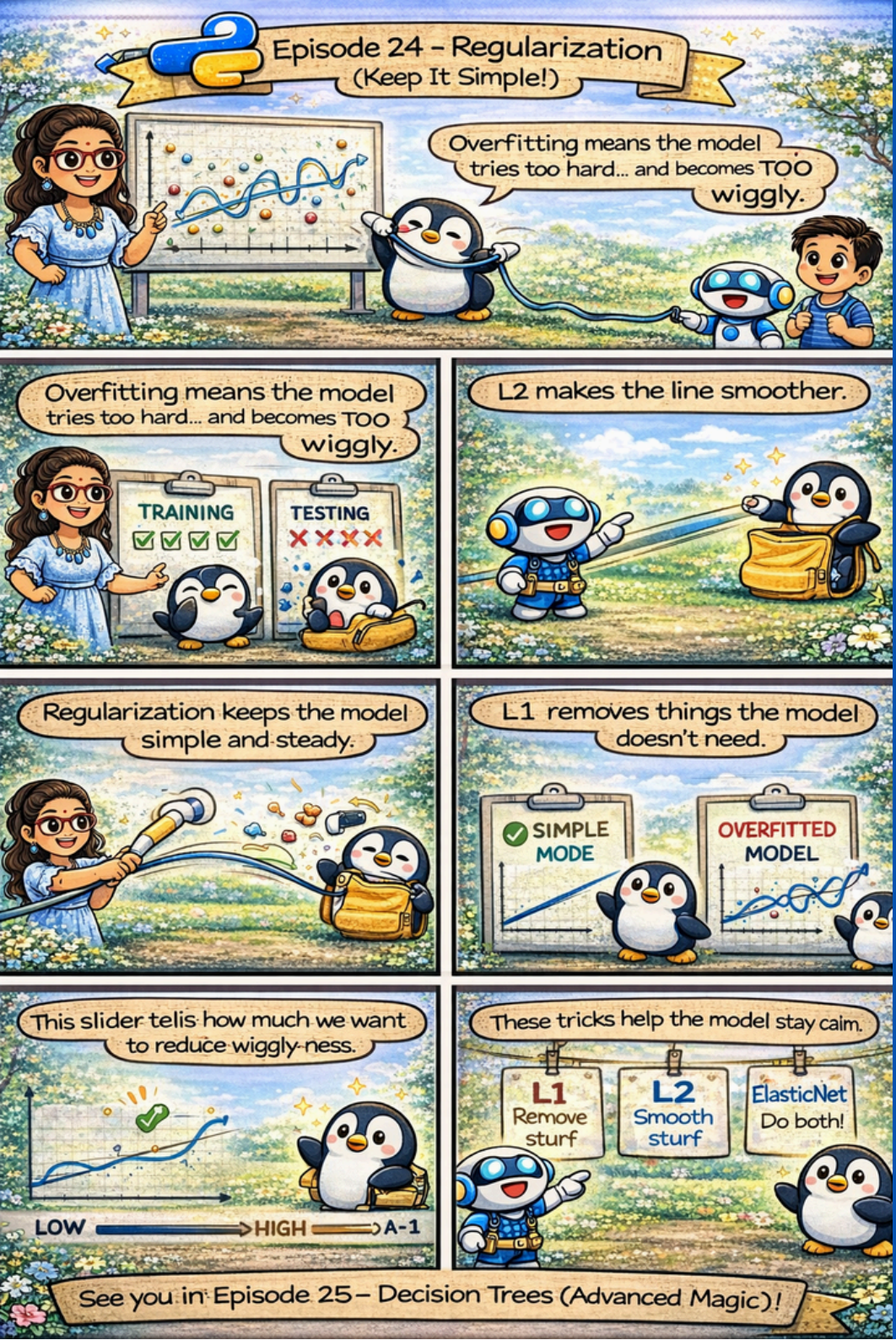

Models need numbers— so we encode categories.
Red
Green
1,0,0
0,1,0
0,0,1

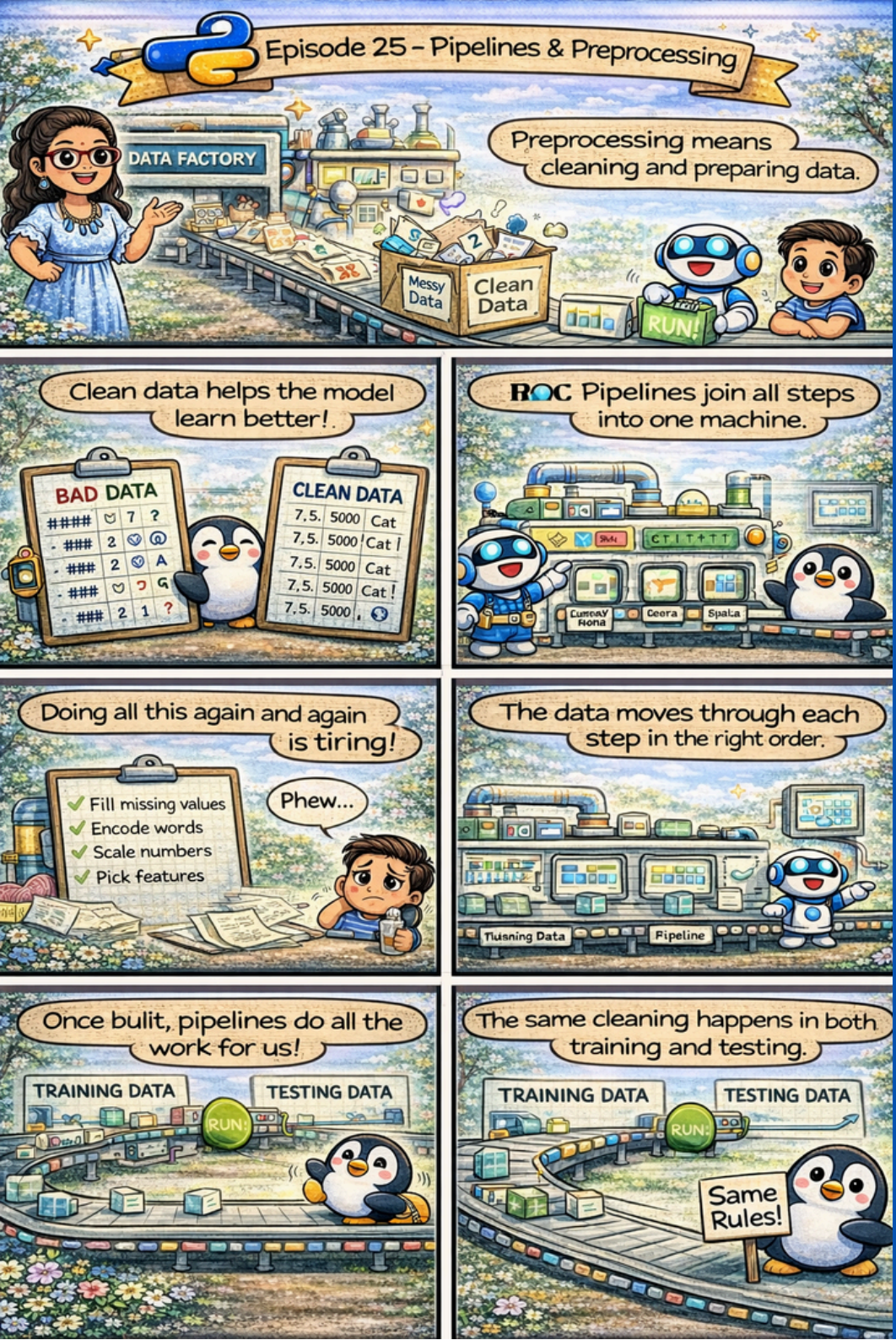

We can group values into buckets!
Child
Teen
Adult
20+
0-12
13-19
13-19
20+
C
Teen
Adult
20+
0-12
13-19
20+
20+

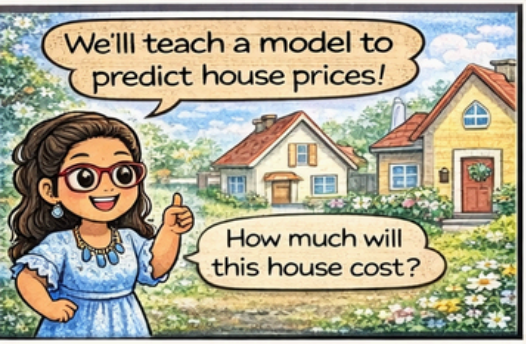

A single column can give many useful features!
Day = 23
Month = 5
Year 2024
Weekday - Thu


Better features = smarter models!
Simple Model
Improved Model
Accuracy
92%
FEATURE TOOLBOX


Episode 23 – Hyperparameter Tuning
These are settings that control how a model learns.
0.01
0.001
SETTINGS
These are settings that control how a model learns.
Learning Rate
0.1
max_depth
0.001
0.01
0.001
But trying values one by one is slow...
0.01
0.01
0.01
0.01
0.001
0.001
Grid Search tests all combinations automatically!
0.1
0.01
0.001
It checks every option, one by one
0.1
0.1.1.0
15.3
3.5,7
With cross-validation, tuning becomes more reliable.
learning rate [0.1, 0.01, 0.001]
max depth [3, 5, 2
neighbors [3, 5, 7
These are the knobs we tune!
learning_rate
max_depth
neighbors
See you in Episode 24 – Regularization (Avoiding Overfitting!

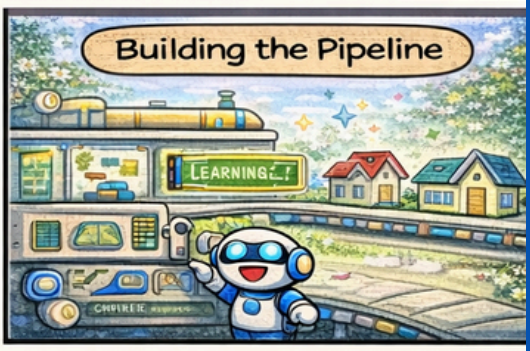

Episode 24 – Regularization
(Keep It Simple!)
Overfitting means the model tries too hard... and becomes TOO wiggly.
Overfitting means the model tries too hard... and becomes TOO wiggly.
TRAINING
TESTING
L2 makes the line smoother.
Regularization keeps the model simple and steady.
L1 removes things the model doesn't need.
SIMPLE MODE
OVERFITTED MODEL
This slider telis how much we want to reduce wiggly-ness.
LOW
HIGH
A-1
These tricks help the model stay calm.
L1
Remove sturf
L2
Smooth sturf
ElasticNet
Do both!
See you in Episode 25 – Decision Trees (Advanced Magic)!

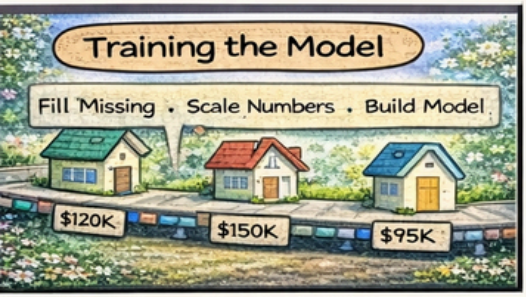

Episode 25 – Pipelines & Preprocessing
DATA FACTORY
Preprocessing means cleaning and preparing data.
Messy Data
Clean Data
RUN!
Clean data helps the model learn better!
BAD DATA
CLEAN DATA
Cat
Cat
Cat
Cat
ROC Pipelines join all steps into one machine.
Ceora
Spata
Doing all this again and again is tiring!
Fill missing values
Encode words
Scale numbers
Pick features
Phew...
The data moves through each step in the right order.
Pipeline
Once bulit, pipelines do all the work for us!
TRAINING DATA
TESTING DATA
RUN:
The same cleaning happens in both training and testing.
TRAINING DATA
TESTING DATA
RUN:
Same Rules!

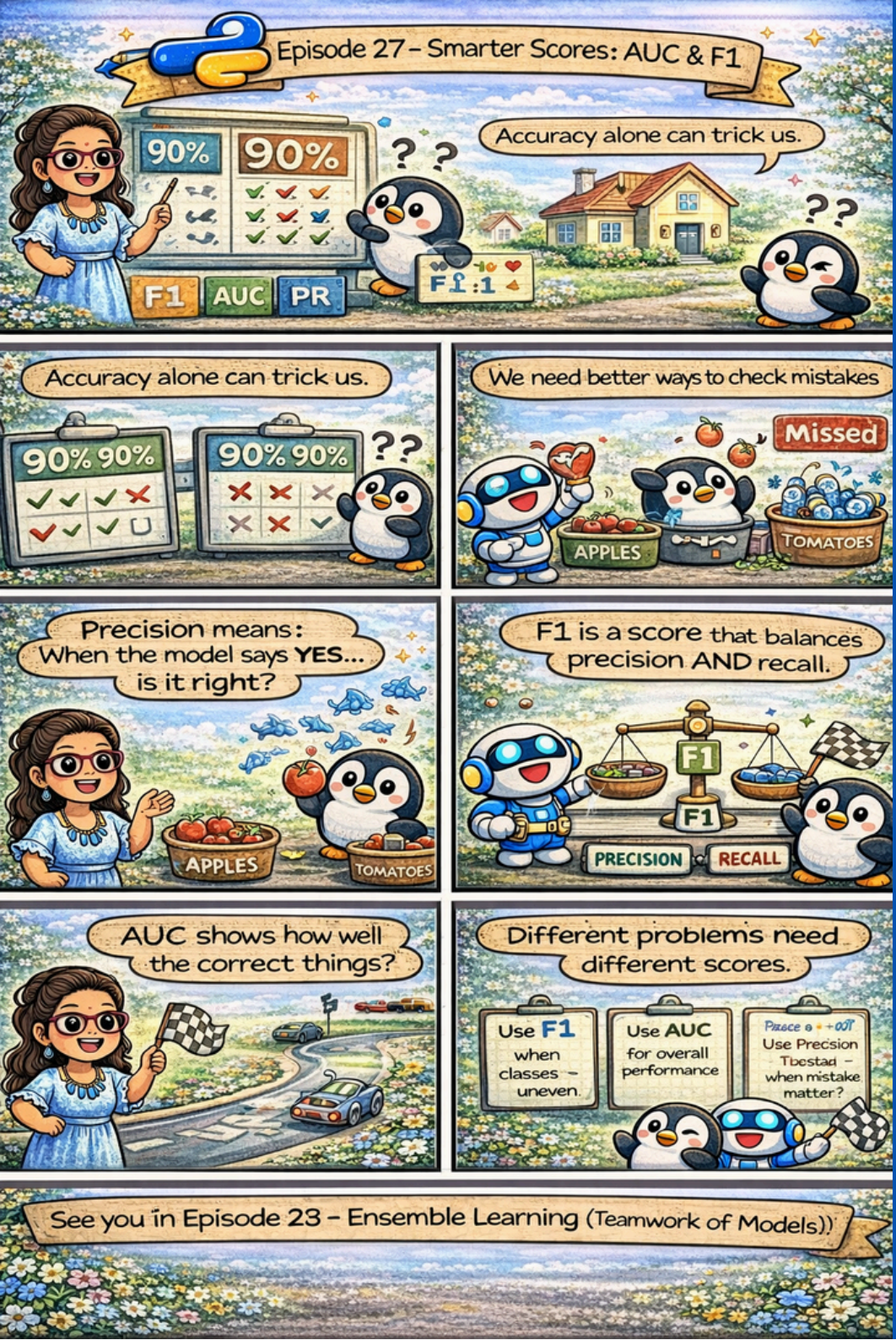


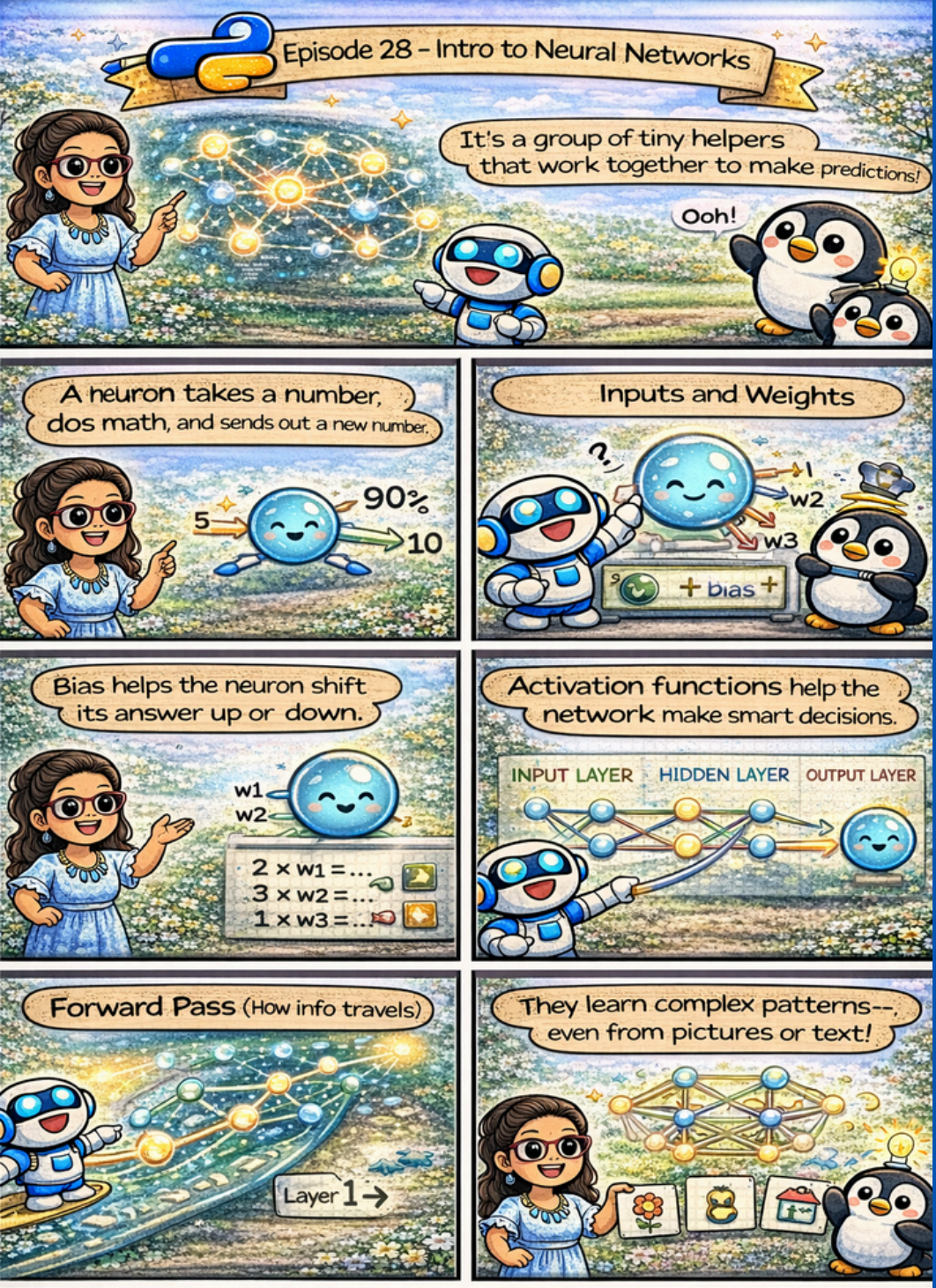
Our Dataset

Size	ISQ FT	8	3	6	✓	Price
700	4,100	6	5	500	5	$1,200
850	6,500	6	3	500	4	$1,500
900	3,500	3	3	Cat	1	$1,500

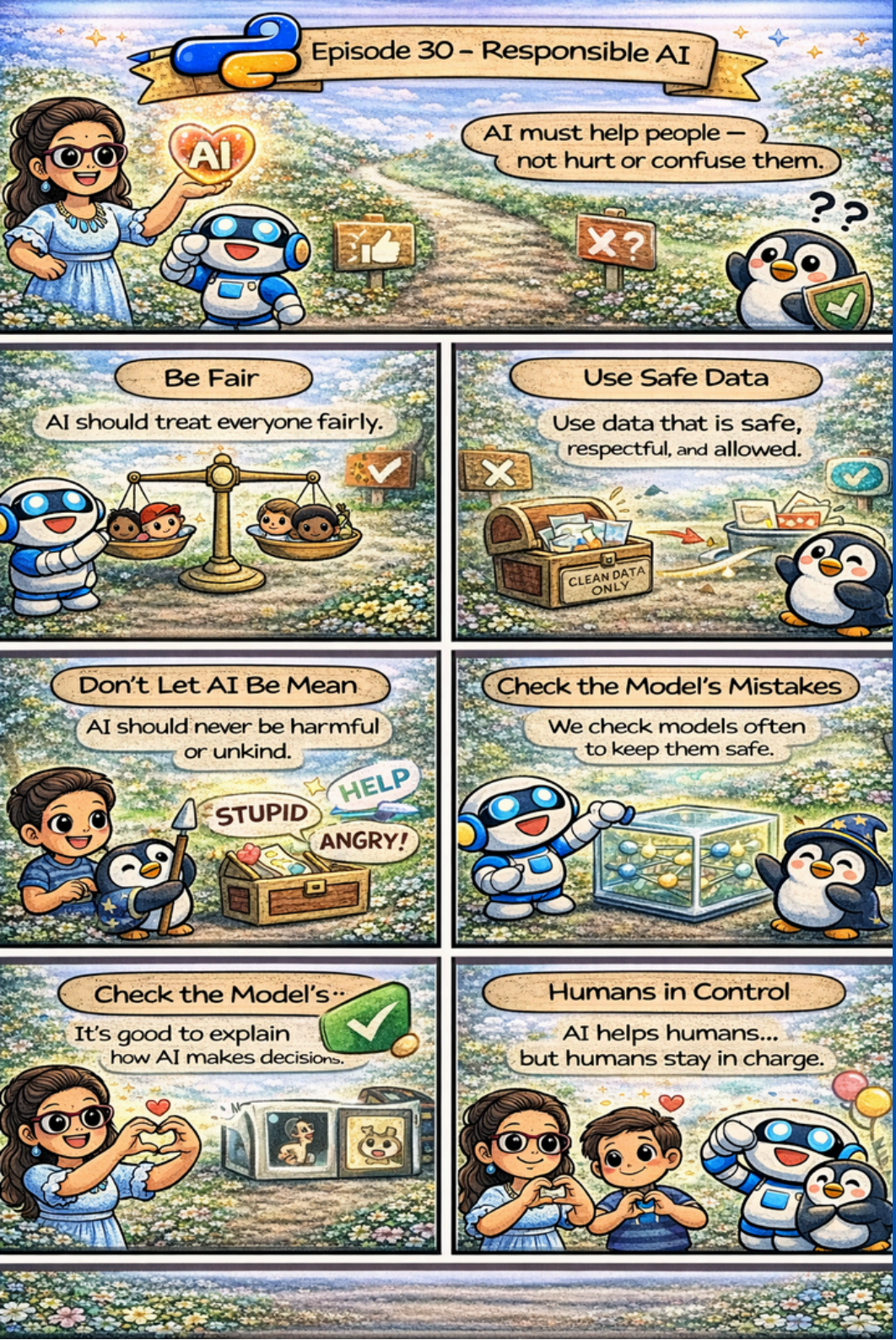





Episode 27 – Smarter Scores: AUC & F1
90%
90%
F1
AUC
PR
Accuracy alone can trick us.
Accuracy alone can trick us.
90% 90%
90% 90%
We need better ways to check mistakes
Missed
APPLES
TOMATOES
Precision means:
When the model says YES...
is it right?
APPLES
TOMATOES
F1 is a score that balances precision AND recall.
F1
F1
PRECISION
RECALL
AUC shows how well the correct things?
Different problems need different scores.
Use F1 when classes – uneven
Use AUC for overall performance
Use Precision when mistake matter?
See you in Episode 23 – Ensemble Learning (Teamwork of Models)!


Episode 28 – Intro to Neural Networks
It's a group of tiny helpers that work together to make predictions!
Ooh!
A neuron takes a number, dos math, and sends out a new number.
5
90%
10
Inputs and Weights
w1
w2
w3
+ bias +
Bias helps the neuron shift its answer up or down.
w1
w2
2 × w1 = ...
3 × w2 = ...
1 × w3 = ...
Activation functions help the network make smart decisions.
INPUT LAYER
HIDDEN LAYER
OUTPUT LAYER
Forward Pass (How info travels)
Layer 1 →
They learn complex patterns--, even from pictures or text!
See you in Episode 29 – Training Neural Networks (Backpropagation Magic)!


Episode 29 – Training Neural Networks (Backpropagation Magic!)
It learns by making guesses... and fixing mistakes!
Backprop Magic
?!
Forward Pass (Making a Guess)
?
The difference between the prediction and the real answer is error.
8
10
Backprop Magic Begins
We send the error backward to fix the network!
10
Fixing Weights Gently
4→2 + 1→3
Smaller weight
Layers Help Each Other
Oops!
I'll fix my part!
I'll adust too!
This happens many times until the network gets really good!
Guess
Error
Backprop
Better Gues
See you in Episode 30 – Mini Project. Build a Neural Network!


Episode 30 – Responsible AI
AI
AI must help people — not hurt or confuse them.
Be Fair
AI should treat everyone fairly.
Use Safe Data
Use data that is safe, respectful, and allowed.
CLEAN DATA ONLY
Don't Let AI Be Mean
AI should never be harmful or unkind.
STUPID
HELP
ANGRY!
Check the Model's Mistakes
We check models often to keep them safe.
Check the Model's··
It's good to explain how AI makes decisions.
Humans in Control
AI helps humans... but humans stay in charge.