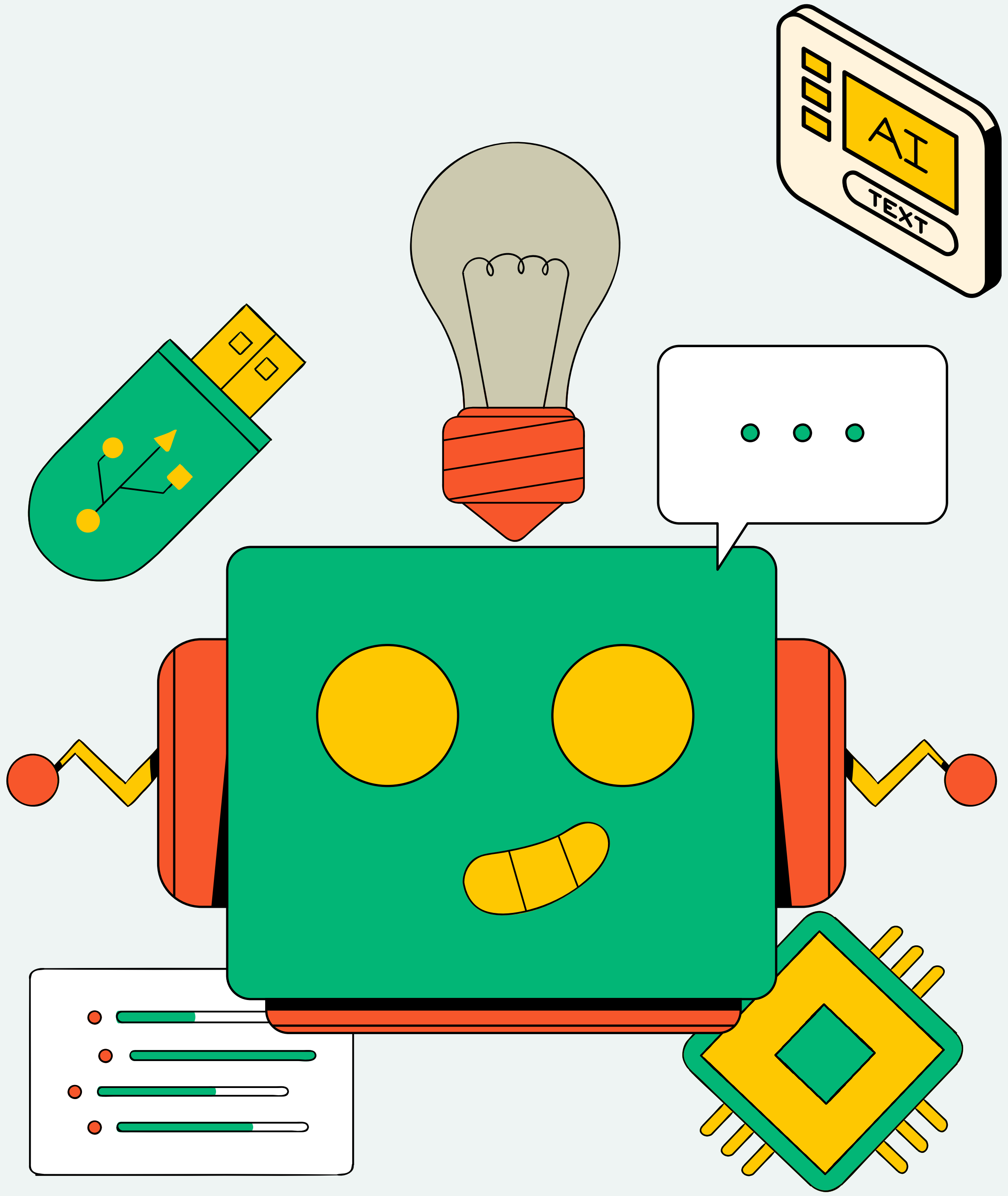
THYNK UNLIMITED

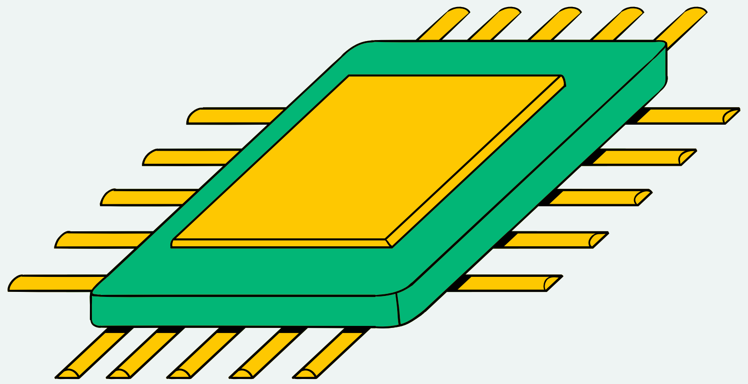
WE LEARN FOR THE FUTURE



ELASTIC NET

ELASTIC NET MODEL – BALANCING L1 & L2 REGULARIZATION

PRESENTED BY: SARAH BELAL



WHAT IS REGULARIZATION?

BEFORE DIVING INTO ELASTIC NET:

01

RIDGE REGRESSION

- Ridge: L2 penalty (squared magnitude of coefficients).
- effective in when features are highly correlated.
- do not performing feature selection

02

LASSO REGRESSION

- Lasso: L1 penalty (absolute magnitude of coefficients).
- not effective when features are highly correlated.
- performing feature selection

03

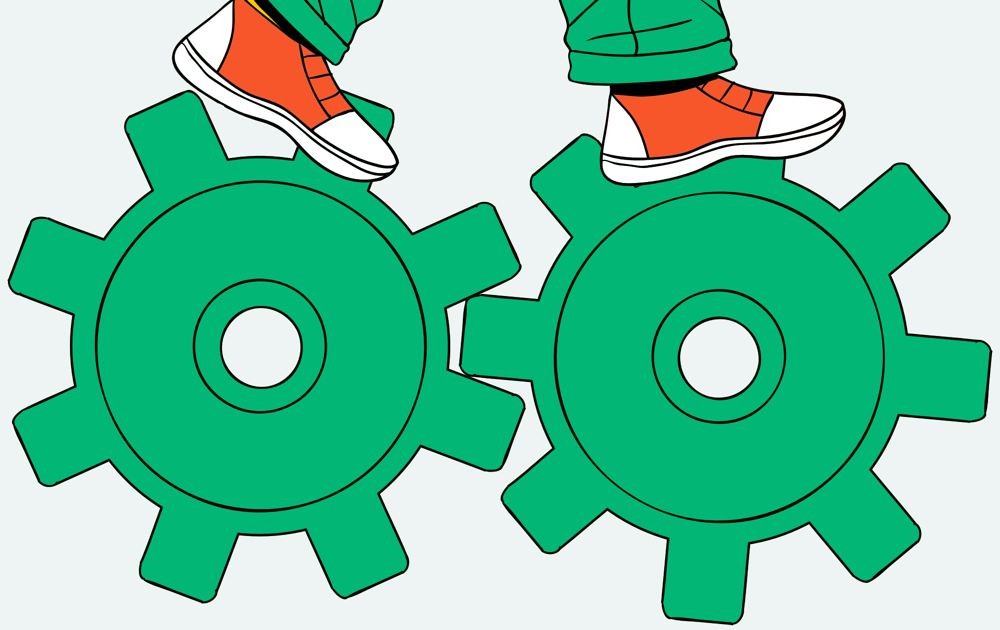
ELASTIC NET

- Elastic Net balances both approaches, keeping groups of correlated features instead of arbitrarily dropping one.
- Balances feature selection (Lasso) and coefficient shrinkage (Ridge)

What is Elastic Net?

Elastic Net is a regularization technique that combines both L1 (Lasso) and L2 (Ridge) regularization. It helps improve model performance by preventing overfitting.

WHY USE ELASTIC NET?



handles high dimensional

In large datasets with thousands of features, Lasso may remove too many, and Ridge may keep too many.

Reduces multicollinearity issues

Elastic Net keeps the important correlated features together.

Prevents over-penalization of correlated features.

Elastic Net ensures important correlated features are preserved without excessive shrinkage.

Useful for feature selection

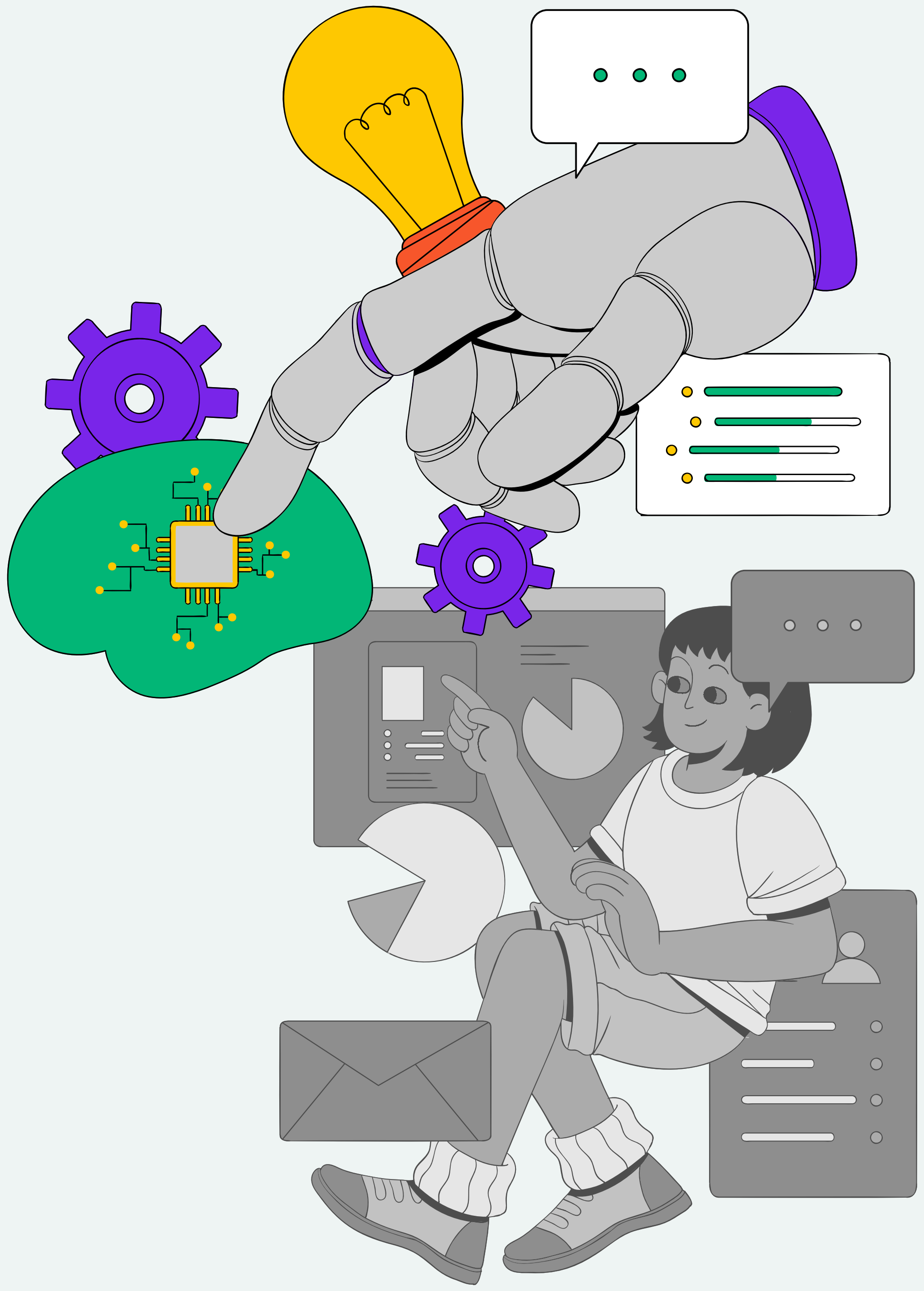
Elastic Net can shrink some coefficients to zero, making it clear which features are most important.

ELASTIC NET FORMULA

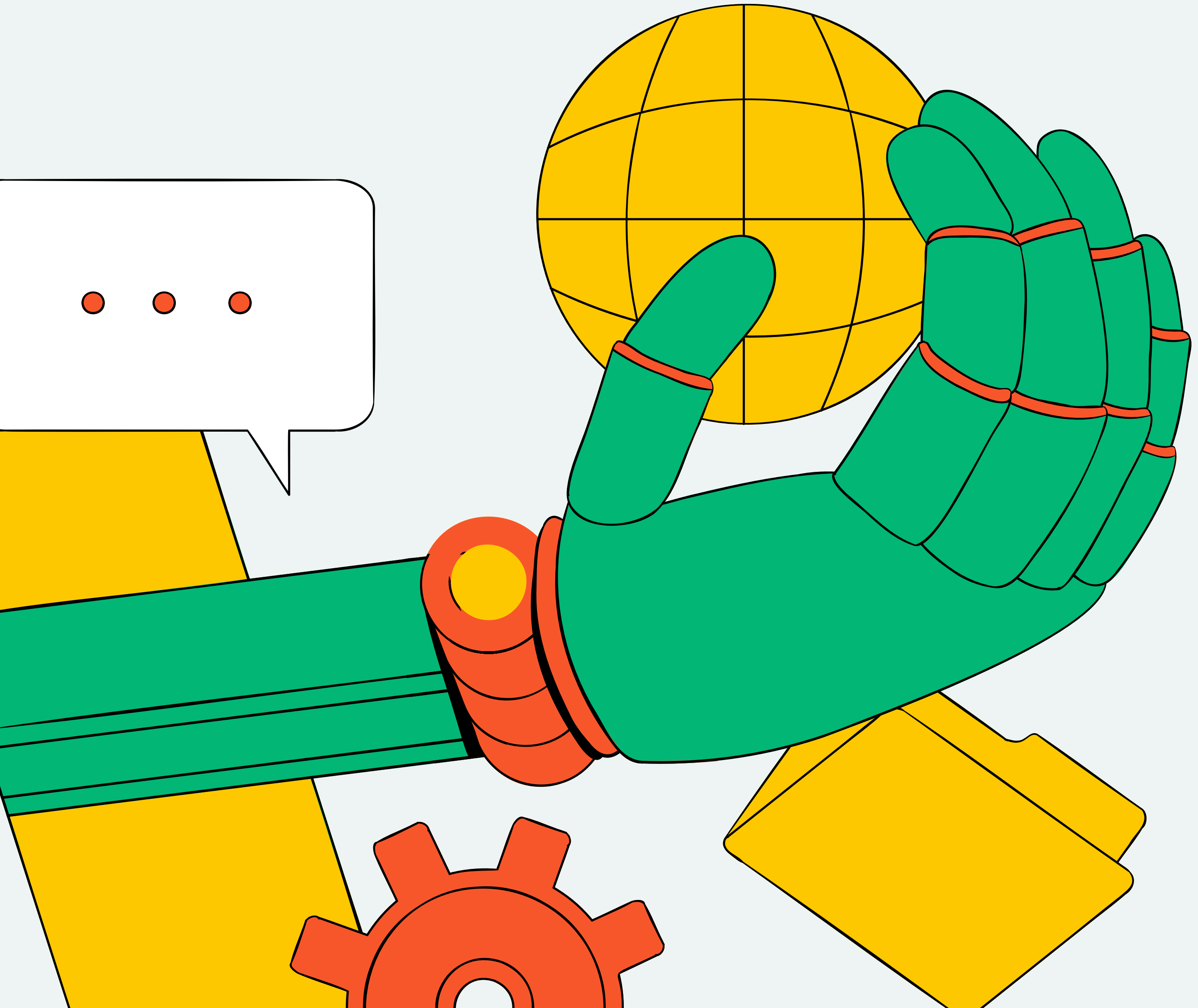
$$\text{Elastic Net Loss} = \frac{1}{2n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \lambda\left(\alpha\sum_{j=1}^{p}|\beta_j| + (1-\alpha)\sum_{j=1}^{p}\beta_j^2\right)$$

Here:

1. $\frac{1}{2n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$: Mean squared error (MSE), the regression loss term measuring the difference between predicted and actual values.
2. λ: Regularization parameter, controlling the overall strength of the regularization.
3. α: Mixing parameter, determining the balance between L1 (Lasso) and L2 (Ridge) penalties.
4. $\sum_{j=1}^{p}|\beta_j|$: L1 penalty, inducing sparsity.
5. $\sum_{j=1}^{p}\beta_j^2$: L2 penalty, encouraging small coefficients.



REAL-WORLD APPLICATIONS



Healthcare & Disease Prediction:

- Used to analyze medical data, selecting the most significant biomarkers for predicting diseases.

THANK YOU