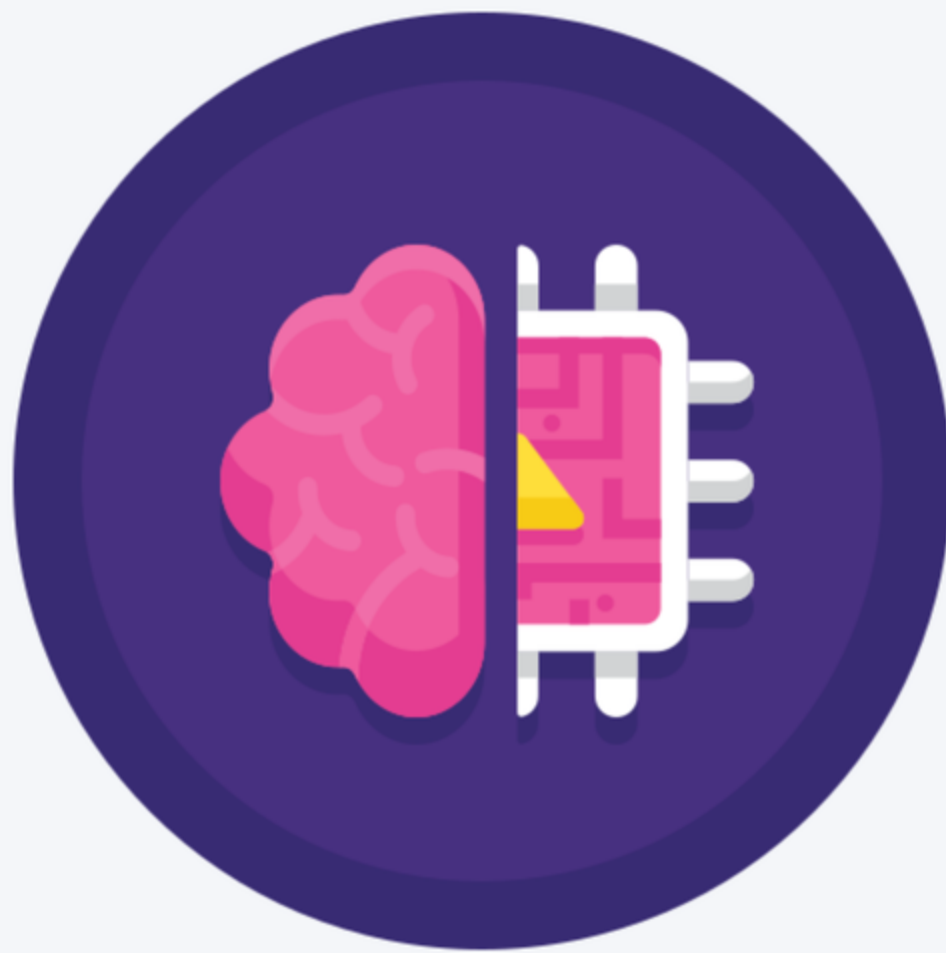


Deconstructing The Machine Learning Magic



Shoaib Akthar
@theshoaibakthar

Follow for more Data Analytics, Data Science and AI insights



The Core Goal

- The system's purpose is to find patterns in data.
- It uses these patterns to make predictions or decisions.
- This is achieved without being explicitly programmed for the task.
- The goal defines what success looks like.
- Common goals include classification, regression, and clustering.



Data Collection

- Raw data is the essential fuel for any ML system.
- This data can be numbers, text, images, or sensor readings.
- It often comes from databases, user interactions, or APIs.
- The quality and quantity of data directly impact results.
- Data must be collected and stored for processing.



Data Preparation

- Raw data is rarely clean and ready to use.
- This step involves cleaning errors and handling missing values.
- Data is often transformed into a consistent numerical format.
- It is split into sets for training, validation, and testing.
- Proper preparation is critical for building an accurate model.



Feature Engineering

- Features are the specific data points the model uses to learn.
- This is the process of selecting and creating the most relevant features.
- Good features help the model identify patterns more easily.
- It can involve combining or modifying existing data columns.
- Effective feature engineering is more art than science.



Model Training

- This is where the algorithm learns from the prepared data.
- The model adjusts its internal parameters to minimize error.
- Training continues until the model's predictions are good enough.
- The training set is used for this learning process.
- The goal is to create a model that generalizes well to new data.



Model Evaluation

- The model must be tested on unseen data to check its performance.
- The holdout test set is used for this final assessment.
- Metrics like accuracy, precision, and recall quantify performance.
- This step ensures the model works in the real world.
- A poor score means going back to improve earlier steps.



Prediction & Inference

- The trained model is now deployed for real-world use.
- It receives new input data and generates predictions or outputs.
- This active use of the model is called inference.
- The system must handle requests quickly and reliably.
- This is the step where the ML system provides its value.



The Feedback Loop

- A crucial component for maintaining a healthy ML system.
- Real-world performance is monitored continuously.
- User feedback and new data are collected.
- This new data is used to retrain and improve the model over time.
- This cycle ensures the model adapts to changing conditions.



Follow for More



Like this post? Repost!

Share your thoughts in the comments!

I share little hacks and tips I actually use in
Data & AI — follow if you wanna see more!



Shoaib Akthar
@theshoaibakthar

Follow for more Data Analytics, Data Science and AI insights