Ad Click-Through Rate Prediction – 1-Minute Recall

Goal:
Predict if a user will click an ad → save money & target users effectively.

Data:

- Features: Daily Time on Site, Age, Area Income, Daily Internet Usage, Gender
- Target: Clicked on Ad (0/1)

Preprocessing:

- Encode Gender: Male=1, Female=0
- Scale features: StandardScaler
- Input shape: 2D → [[...]]

Model:

- Logistic Regression (binary classification)
- Predicts probability of click (model.predict_proba)

Prediction Function:

```
def predict_ad_click(Daily_Time, Age, Income, Internet, Gender):
  gender = 1 if Gender.lower()=='male' else 0
  features = np.array([[Daily_Time, Age, Income, Internet, gender]])
  features_scaled = scaler.transform(features)
  return model.predict_proba(features_scaled)[0][1]*100
```

Challenges & Solutions:

1. Extra country features → removed for simplicity
2. Scaling mismatch → applied StandardScaler
3. Encoding categorical features → simple 0/1
4. Shape mismatch → ensured 2D array input

Deployment:

- Streamlit app → input features → output click probability
- Joblib used to load model & scaler
- @st.cache_resource for efficiency

Key Takeaways:

- Simple model + proper preprocessing → accurate & interpretable
- Logistic Regression good for small datasets & binary targets

- Probability output allows ranking users by likelihood to click

Quick Memory Map:
Input → Encode & Scale → Logistic Regression → Predict Probability → Output