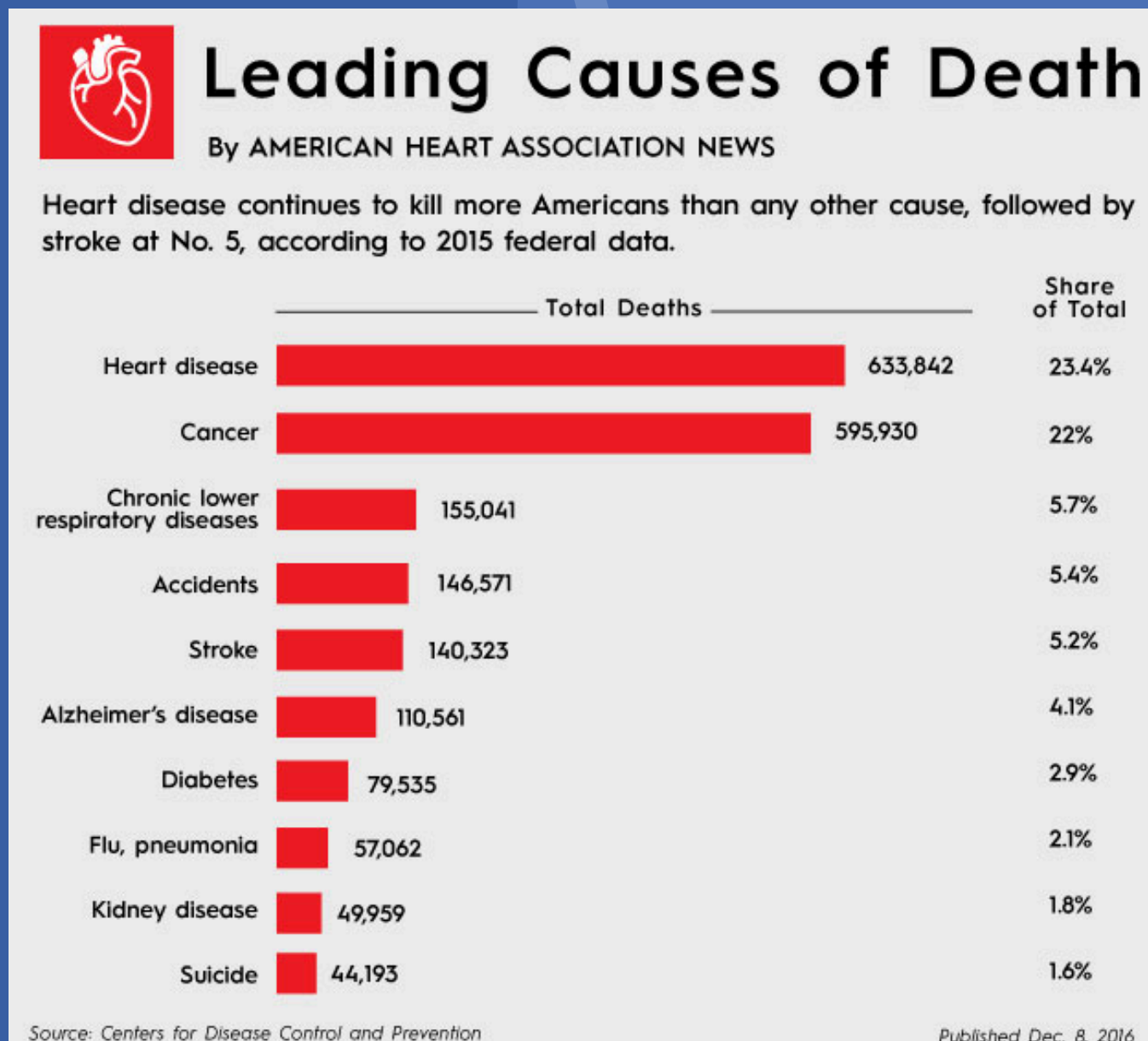




# PREVENTI-SCAN

AI Heart Disease Detector

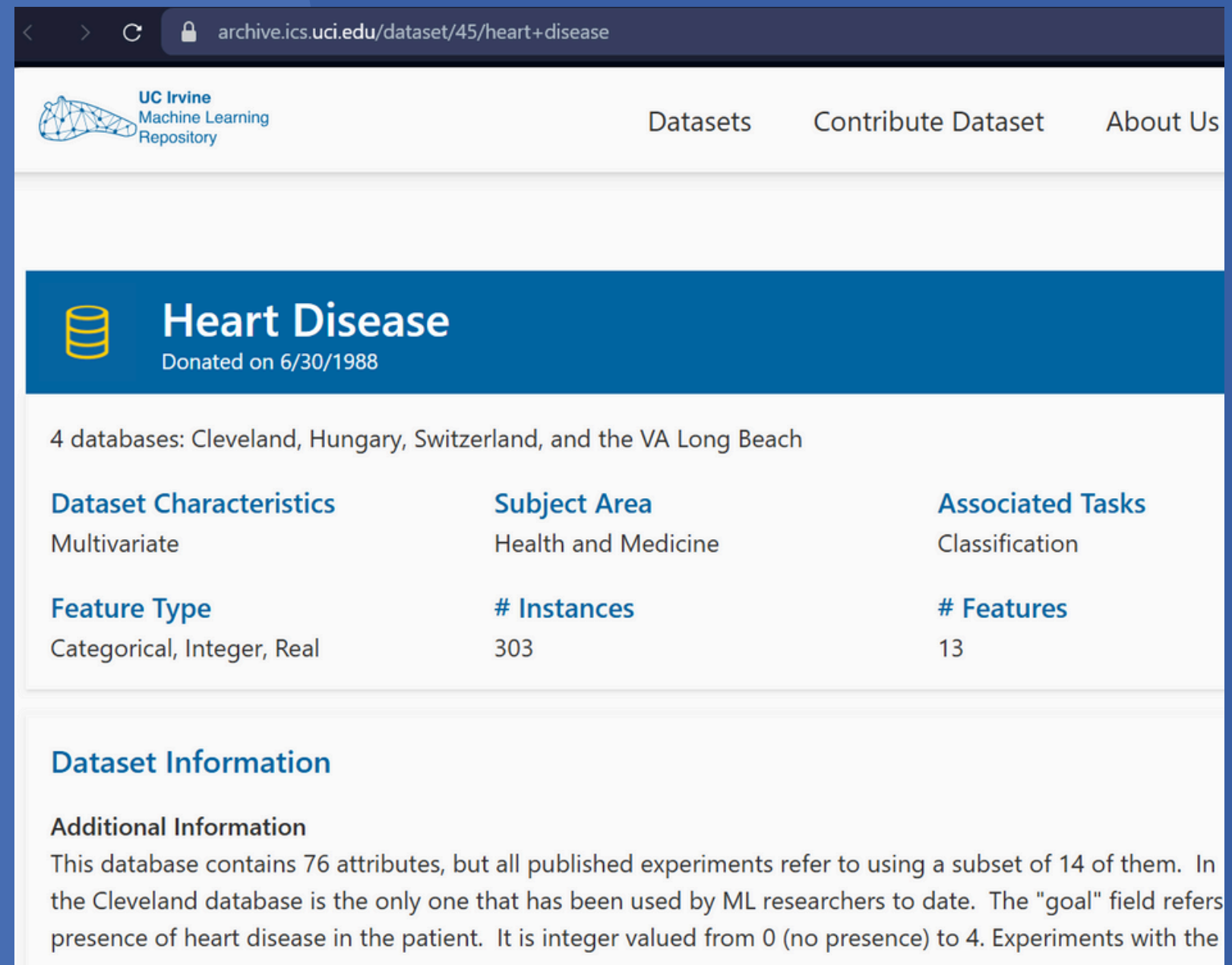
# The Vision & Purpose



- Heart disease is the #1 cause of death in the world.
- Many people don't know they are at risk.
- Doctors don't have time to check every patient manually.
- I built an AI to predict risk from patient data in seconds.
- This can help save lives with faster and smarter decisions.

# How I Prepared the Data

- Data from real patients at Cleveland Clinic (Converted to CSV file)
- Total of 297 patient records
- 13 health features like age, cholesterol, heart rate, chest pain
- Cleaned using Python and pandas
- Target column shows if a person has risk (1) or no risk (0)



The screenshot shows the UC Irvine Machine Learning Repository page for the Heart Disease dataset. The page is titled "Heart Disease" and indicates it was donated on 6/30/1988. It lists 4 databases: Cleveland, Hungary, Switzerland, and the VA Long Beach. The page includes a table with dataset characteristics, subject area, and associated tasks. Below the table, there is a section for dataset information, including additional information about the dataset's attributes and usage.

Dataset Characteristics	Subject Area	Associated Tasks
Multivariate	Health and Medicine	Classification

Feature Type	# Instances	# Features
Categorical, Integer, Real	303	13

**Dataset Information**

**Additional Information**

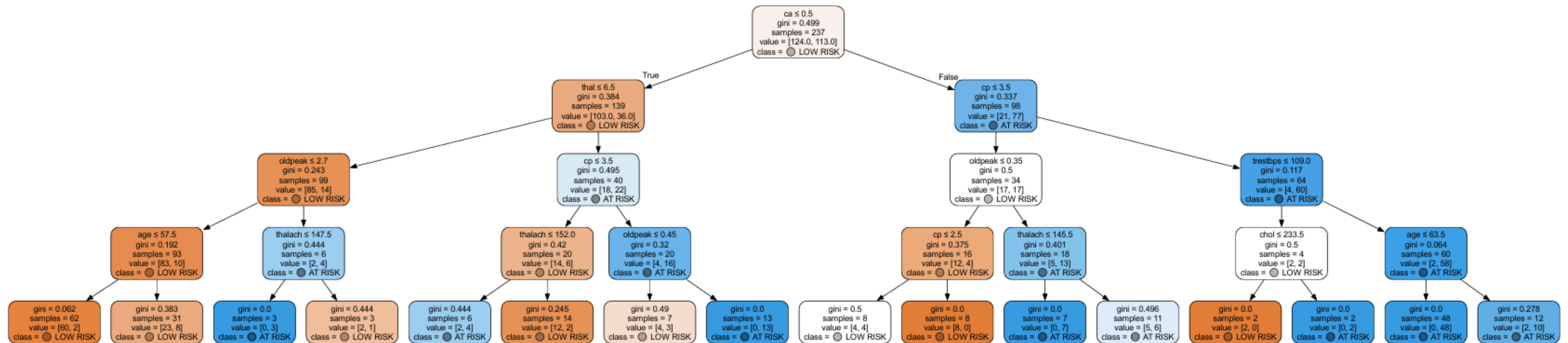
This database contains 76 attributes, but all published experiments refer to using a subset of 14 of them. In the Cleveland database is the only one that has been used by ML researchers to date. The "goal" field refers to presence of heart disease in the patient. It is integer valued from 0 (no presence) to 4. Experiments with the

# How the AI Thinks: The Decision Tree

- The decision tree is trained using real patient data
- It learns to ask yes/no health questions like “Is cholesterol > 240?”
- Each branch splits based on answers to these questions
- At the end of each path, it gives a result:  
● AT RISK or ● LOW RISK
- This part of the code creates and draws the tree using scikit-learn + Graphviz

```
35
36 # ✅ Step 8: Train decision tree
37 clf = DecisionTreeClassifier(max_depth=4)
38 clf.fit(X_train, y_train)
39
40 # ✅ Step 9: Visualize tree
41 dot_data = export_graphviz(
42     clf,
43     out_file=None,
44     feature_names=X.columns,
45     class_names=["● LOW RISK", "● AT RISK"],
46     filled=True,
47     rounded=True,
48     special_characters=True
49 )
50 graph = graphviz.Source(dot_data)
51 graph.render("PreventiScan_Heart_Tree")
52 graph.view()
53
54 # ✅ Step 10: Predict patient risk
55 df['PredictedRisk'] = clf.predict(X)
56 df['PredictedRisk'] = df['PredictedRisk'].apply(lambda x: "● AT RISK" if x == 1 else "● LOW RISK")
57
58 # ✅ Step 11: Show risk summary
59 total = len(df)
60 at_risk = df[df['PredictedRisk'] == "● AT RISK"]
61 low_risk = df[df['PredictedRisk'] == "● LOW RISK"]
```

# Decision Tree



- The AI moves through the tree until it finds a final result.

● Blue boxes = At Risk

● Orange boxes = Low Risk



# Risk Prediction Results

- The AI analyzed 297 patients from the dataset
- It found 131 patients At Risk and 166 Low Risk
- Each patient got a risk label using medical data
- This helps doctors see who needs urgent care first
- The system is fast, smart, and can scale to millions

PREVENTISCAN HEART DISEASE RISK SUMMARY						
Total Patients Analyzed: 297						
At Risk Patients : 131						
Low Risk Patients : 166						
FIRST 10 PATIENT RISK PREDICTIONS:						
	PatientID	age	chol	trestbps	thalach	PredictedRisk
0	1	63.0	233.0	145.0	150.0	LOW RISK
1	2	67.0	286.0	160.0	108.0	AT RISK
2	3	67.0	229.0	120.0	129.0	AT RISK
3	4	37.0	250.0	130.0	187.0	LOW RISK
4	5	41.0	204.0	130.0	172.0	LOW RISK
5	6	56.0	236.0	120.0	178.0	LOW RISK
6	7	62.0	268.0	140.0	160.0	AT RISK
7	8	57.0	354.0	120.0	163.0	LOW RISK
8	9	63.0	254.0	130.0	147.0	AT RISK
9	10	53.0	203.0	140.0	155.0	AT RISK
10	11	57.0	192.0	140.0	148.0	LOW RISK
11	12	56.0	294.0	140.0	153.0	LOW RISK
12	13	56.0	256.0	130.0	142.0	AT RISK
13	14	44.0	263.0	120.0	173.0	LOW RISK
14	15	52.0	199.0	172.0	162.0	LOW RISK
15	16	57.0	168.0	150.0	174.0	LOW RISK
16	17	48.0	229.0	110.0	168.0	LOW RISK
17	18	54.0	239.0	140.0	160.0	LOW RISK
18	19	48.0	275.0	130.0	139.0	LOW RISK
19	20	49.0	266.0	130.0	171.0	LOW RISK
20	21	64.0	211.0	110.0	144.0	LOW RISK
21	22	58.0	283.0	150.0	162.0	LOW RISK
22	23	58.0	284.0	120.0	160.0	LOW RISK

# Why This Matters

- This AI can help doctors save time and lives
- Can be used in hospitals, clinics, or mobile apps
- Gives instant health risk feedback from raw data
- Can be trained on more diseases in the future



# References & Resources

Dataset Source:

UCI Heart Disease Dataset.

Python Libraries Used:

pandas, scikit-learn, graphviz.

Data Cleaning:

Used pandas to clean and prepare real patient records.

Model Used:

Built a DecisionTreeClassifier with scikit-learn.

Code Developed By:

Faiaz Ahmed, PreventiScan AI Project (2025).





THANK  
YOU ♥ +

