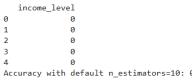
LAB 8 ADA BOOST:

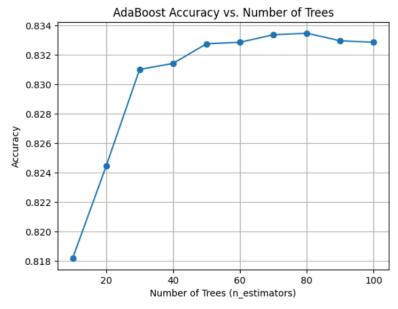
```
In [2]:
         #income
         import pandas as pd
         from sklearn.model_selection import train_test_split
         from sklearn.ensemble import AdaBoostClassifier
         from sklearn.tree import DecisionTreeClassifier
         from sklearn.metrics import accuracy_score
         import matplotlib.pyplot as plt
         from google.colab import files
         # Upload the CSV file
         uploaded = files.upload()
         # Read the CSV file into a pandas DataFrame
         df = pd.read_csv(next(iter(uploaded))) # Load the first uploaded file
         print(df.head())
         # Split features and target
         X = df.drop("income_level", axis=1)
         y = df["income_level"]
         # Encode target if necessary (e.g., if it's strings)
         y = y.astype('category').cat.codes # Converts labels to numeric
         # Train-test split
         X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
         # 1. Train AdaBoost with default n_estimators = 10
         ada default = AdaBoostClassifier(
             estimator=DecisionTreeClassifier(max_depth=1),
             n_estimators=10,
             random_state=42
         ada_default.fit(X_train, y_train)
         y_pred_default = ada_default.predict(X_test)
         default_score = accuracy_score(y_test, y_pred_default)
         print(f"Accuracy with default n_estimators=10: {default_score:.4f}")
       # 2. Fine-tune n_estimators
       scores = \{\}
       for n in range(10, 110, 10): # Try 10 to 100 in steps of 10
          ada = AdaBoostClassifier(
              estimator=DecisionTreeClassifier(max_depth=1),
              n_estimators=n,
              random_state=42
          ada.fit(X_train, y_train)
          y_pred = ada.predict(X_test)
           acc = accuracy_score(y_test, y_pred)
           scores[n] = acc
       # Find best score and corresponding number of estimators
       best_n = max(scores, key=scores.get)
       best_score = scores[best_n]
       print(f"Best accuracy: {best_score:.4f} with n_estimators={best_n}")
       plt.plot(list(scores.keys()),\ list(scores.values()),\ marker='o')
       plt.xlabel("Number of Trees (n_estimators)")
       plt.ylabel("Accuracy")
       plt.title("AdaBoost Accuracy vs. Number of Trees")
       plt.grid(True)
       plt.show()
```

Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

Saving income.csv to income (1).csv age fnlwgt education_num capital_gain capital_loss hours_per_week \ 13 2174 0 39 1 50 0 77516 0 40 0 83311 13 0 13 2 38 215646 3 53 234721 9 7 0 0 40 0 0 40 4 28 338409 13 0 0 40



Accuracy with default n_estimators=10: 0.8182 Best accuracy: 0.8335 with n_estimators=80



				(18	D					
5/5/202 Mondo		ab-8		of Deeler						
Shoral	DINO B	- 400	Don	tons						
	Activo	Ada Boost - tot. Ada Boost - tot. Job profile yes -5 +1, else - 1 Job profile (GPA JOB POWELLE Label)								
	*05 PM	ofile yes	Tob Powfile	label	1					
	ander		yes	+1	1					
	1	29	yes	+1						
11 2021	2	29	NO	-1						
	3	29	NO	-1						
	4	29	yes	+1	AC					
	5	29	yes yes	+1	1					
	6 29									
	PS	1 1 : /		1						
,	Creade	dala 3/	Tour label.	Poudi	tion					
0	Endex	CAPA	+1	+1	~					
1605	offer Krowll	0	+1	-1	×					
	2	, N	mar.9 -1-	+1	×					
· De	3 1 NOX		1	-1	~					
	4	0	2	+1	V					
OCA= KSA		S Molege	+)	+1	V					
	6		.,							
	Mark Hall State				-					
Part	If CasAs	9 - Pred	ict + 1, Elec	- J 01	The St					
modern	influence	a hu	cy score	neusia						
The state of	Each weigh	6ht Levilia	lly = 1/6 =	0.1667						
	Encoure	ct and	ex - 2,3	anishi.						
	Everor wight = D. + D. = D. 1167 + 5 1117									
	Escror wight = $D_1 + D_3 = 0.1667 + 0.1667$ $= 0.3333$									
		01-10			1					
	d= ln	(1-canox)	= 1 ln (1 - 0, \begin{aligned} 2 \(\text{0.} \delta 3333 \\ \text{3.} \\ \text{2.} \(\text{0.} \delta 3333 \\ \text{2.} \\ \text{2.} \(\text{0.} \delta 3333 \\ \text{2.} \\ \text{2.} \(\text{2.} \delta 3333 \\ \text{2.}	Meuro	-					
	2 (ewor)	2 1 1 1 - 0,333	3) = 0.3	466					
	if capa	29	0.8333							
	3.	les Per	edict job pa	011.	V.0					
	9 tup	mp 100	edset job por		100,					
A WOOD SHOT		weeph	= 0.31	,						
1	Alle I	Bally No	46	6						

0			9	Date Page 19			
ROY	For second datalet,						
40	best accuracy score = 0.8335						
	Number	of traces =	80				
	1. 1. 5 5 6	A STATE	(3)	2.70			
	Confueio	materia.					
	3 9	2 2 10	2 8	284			
	=50X	7 13 0	8				
	> 50 K	1343	3.5	1012			
sua	- 30 2	13	43	9 11			
		2=501	c	>50K			
	Service Street		redicte	d.			
		(3)	12 20	charles East			
Sh	1			inches f			
25	125	(5,7) chutes	Lister 4	Palagoint C			
5		ET.	Ca	4.1			
		137:25	1195	15,2			
	LATE	113	413	13.5			
	e	0	153	. 63			
	c	Ecol .	2006	85,5			
	5	Just .	13838	4.5,5			
		Jack	125	2545			
		6	Chickon	Triples +			
	Per State of the S		631	(11)			
Tage Fo		()	62.2	(1.5,2)			
			Cus	(34)			
			(3.5 W				
		200	202)				
	THE PAPER	(2) . (2	4 4 5 5	Avrage (16			