

# Sperimentazione in WEKA

Confronto tra classificatori: Naive Bayes, Decision Tree e Random Forest

ESAME DI INTELLIGENZA ARTIFICIALE (9CFU)

PROF. STEFANO FERILLI

Davide Scaraggi – 661635

## Sommario

1	Obiettivi della sperimentazione .....	3
2	Algoritmi di apprendimento.....	4
2.1	Classificatori generali e classificatori discriminativi .....	4
2.2	Naive Bayes.....	5
2.3	Decision tree.....	5
2.4	Random forest .....	6
3	Dataset utilizzati.....	8
3.1	Scelta dei dataset .....	8
3.2	Adult dataset.....	8
3.3	Vehicle Silhouettes dataset.....	9
4	Sperimentazione.....	10
4.1	Fase preliminare .....	10
4.2	Sperimentazione con parametri di default.....	10
4.2.1	Adult dataset .....	10
4.2.2	Silhouettes dataset.....	18
4.3	Sperimentazione con parametrizzazione.....	23
4.3.1	Tuning dei parametri.....	23
4.3.2	Silhouettes dataset.....	25
4.4	Test di valutazione .....	25
5	Modelli generati .....	28
5.1	Modello di Adult con il Naive Bayes.....	28
5.2	Modello di Adult con il Decision Tree.....	32
5.3	Modello di Silhouette con il Naive Bayes.....	76
5.4	Modello di Decision Tree su Silhouette.....	79

## 1 Obiettivi della sperimentazione

La seguente sperimentazione è mirata alla valutazione delle prestazioni di alcuni algoritmi di apprendimento automatico applicati a task di **classificazione**. La sperimentazione viene eseguita utilizzando dataset differenti sia in termini di contesto applicativo che di numero di istanze, di attributi e di classi. L'obiettivo è quello di mettere in evidenza i punti di forza e di debolezza di ciascun algoritmo. Risulta indispensabile eseguire la sperimentazione su varie tipologie di dataset in quanto, anche se le performance di un modello dovessero dipendere a priori dal numero di features, dal numero delle osservazioni e dal tipo di osservazioni, non è ovvio che un modello abbia performance migliori rispetto ad un altro. La ragione di quanto appena detto è che, ad esempio, è molto importante considerare anche il tempo necessario per l'apprendimento del modello. Le performance dell'algoritmo di apprendimento possono essere fortemente influenzate anche dai parametri di configurazione.

Nella scelta dell'algoritmo di apprendimento più opportuno, il **tradeoff Bias-Variance** ricopre un ruolo determinante. Il *tradeoff Bias-Variance* si ricollega al principio del **rasoio di Occam**, secondo cui tra le soluzioni possibili ad una problematica va selezionata quella che effettua meno assunzioni. Per **bias** s'intende la differenza tra le predizioni attese dal modello e i valori reali. Infatti, alcuni algoritmi risultano troppo rigidi per apprendere il contenuto di un dataset. Il concetto di **varianza**, invece, fa riferimento alla sensibilità dell'algoritmo verso specifici insiemi di dati di training. Algoritmi con alta varianza producono modelli fortemente differenti a seconda del training set utilizzato. Algoritmi caratterizzati da **alto bias e bassa varianza** apprendono modelli più consistenti ma meno accurati e la bassa varianza è dovuta ad un algoritmo meno complesso, con struttura sottostante rigida (ad es. Naive Bayes, Regression, ...). Algoritmi caratterizzati da **alta varianza e basso bias** apprendono modelli più accurati ma meno consistenti e il basso bias è dovuto ad un algoritmo più complesso, con una struttura sottostante flessibile (Decision Trees, ...). Un tradeoff tra bias e varianza è necessario perché:

- Algoritmi con una struttura molto semplice (alto bias e bassa varianza) comportano il fenomeno dell'**underfit**, ovvero modelli che non riescono a catturare del tutto il contenuto del dataset di training;
- Algoritmi con una struttura molto complessa (alta varianza e basso bias) comportano il fenomeno dell'**overfit**, ovvero modelli che memorizzano molto rumore presente nei dati piuttosto che i segnali fondamentali per la classificazione.

Dinanzi ad un training set di piccole dimensioni, i classificatori con alto bias e bassa varianza si comportano meglio rispetto ai classificatori con alta varianza e basso bias in quanto questi ultimi portano al fenomeno dell'overfitting. Tuttavia, i classificatori con alta varianza e basso bias risultano l'alternativa

migliore con il crescere del training set, dato che i classificatori con alto bias risultano meno performanti con dataset di grandi dimensioni. Nella seguente sperimentazione sono stati considerati i seguenti classificatori:

- **Naive Bayes**, per via della sua semplicità;
- **Decision Trees**, come modello più complesso rispetto al Naive Bayes;
- **Random Forests**, per rimediare all'overfitting generato dai Decision Trees.

## 2 Algoritmi di apprendimento

### 2.1 Classificatori generali e classificatori discriminativi

Nelle sperimentazioni che verranno descritte di seguito, sono stati utilizzati diversi metodi di classificazione. Sono distinguibili due categorie di classificatori:

- **generativi**
- **discriminativi**

Considerati dati di input  $X$  da classificare nelle labels  $Y$ , un modello generativo studia la distribuzione di probabilità congiunta  $P(x,y)$  e un modello discriminativo studia la distribuzione di probabilità condizionata  $P(y|x)$ . La distribuzione di probabilità condizionata  $P(y|x)$  è la distribuzione naturale per classificare un dato esempio  $x$  in una classe  $y$ , motivo per cui gli algoritmi che la utilizzano sono chiamati discriminativi. Gli algoritmi generativi, invece, modellano  $P(x,y)$ , che può essere trasformata in  $P(y|x)$  applicando il teorema di Bayes.

$$P(y|x) = \frac{P(x|y) P(y)}{P(x)}$$

**Con un classificatore generativo:** si cerca di apprendere il modello che genera i dati effettuando delle assunzioni sulle distribuzioni di probabilità. Si assume che il modello appreso catturi quello reale. Un esempio è rappresentato dal classificatore *Naive Bayes*: si stimano  $P(x|y)$  e  $P(y)$  dai dati, con  $P(x)$  costante, e successivamente si moltiplica  $P(x|y)P(y)$  secondo il teorema di Bayes:

**Con un classificatore discriminativo:** si cerca di apprendere il modello basandosi esclusivamente sui dati osservati, effettuando poche assunzioni sulle distribuzioni di probabilità. In questo caso, i risultati dipendono in maniera rilevante dalla qualità dei dati, che devono essere abbastanza e soprattutto rappresentativi. Un esempio è rappresentato dalla *Logistic Regression*:  $P(y|x)$  viene stimato direttamente dai dati.

Generalmente, i modelli discriminativi, rispetto a quelli generativi, consentono il raggiungimento di

risultati migliori nei task di classificazione se hanno in pasto abbastanza dati. I classificatori generativi studiano  $P(y|x)$  indirettamente, col rischio di effettuare assunzioni sbagliate sulle distribuzioni di probabilità. Vladimir Vapnic, nel suo libro “*Statistical Learning Theory*”, afferma che un problema di classificazione andrebbe risolto direttamente e mai risolvendo un problema più generale come passo intermedio (ad esempio,  $P(x|y)$ ). Andrew Ng, nel paper “*On Discriminative vs Generative classifiers*”, afferma che il modello generativo porta ad un errore asintotico (al crescere del numero degli esempi di training) più grande rispetto al modello discriminativo.

## 2.2 Naive Bayes

Il classificatore di Naive Bayes si basa sul teorema di Bayes con l’assunzione di indipendenza tra gli attributi considerati. In particolare, assume che la presenza di un particolare attributo in una classe sia indipendente dalla presenza di qualsiasi altro attributo. Se la condizione di indipendenza è verificata, il Naive Bayes sarà in grado di convergere più velocemente di un modello discriminativo come la Logistic Regression, utilizzando una quantità inferiore di dati. Vi sono alcuni casi, in cui, anche in presenza di dipendenza tra gli attributi l’algoritmo è in grado di raggiungere dei buoni risultati di classificazione.

I **vantaggi** principali sono i seguenti:

- Facile da implementare
- Efficiente in tempi di esecuzione e in spazio richiesto
- Lavora bene anche con piccole quantità di dati, rispetto ai *Decision Tree*
- Minor tendenza all’*overfitting*
- Sotto specifiche condizioni può fornire risultati migliori rispetto a metodi più complessi.

Gli **svantaggi** principali sono i seguenti:

- Effettua assunzioni, assumendo l’indipendenza tra gli attributi
- Rischio di underfitting

## 2.3 Decision tree

I Decision Tree implementano modelli di classificazione oppure di regressione utilizzando una struttura ad albero. Il dataset viene diviso in sottoinsiemi man mano più piccoli e allo stesso tempo viene incrementalmente sviluppato un albero di decisione, formato di nodi di decisione e foglie. Un nodo di decisione fornisce due o più alternative di scelta e le foglie rappresentano il risultato di classificazione oppure la decisione. Quindi, un albero di decisione risponde a domande in sequenza che portano a seguire un percorso specifico che si conclude con la classificazione. Il concetto di profondità risulta fondamentale perché indica il numero di domande a cui è necessario dare risposta per raggiungere la

classificazione.

I **vantaggi** principali sono i seguenti:

- I risultati sono di facile interpretazione e rappresentazione
- I meccanismi interni possono essere osservati e questo rende possibile riprodurre il lavoro effettuato e verificare i risultati ottenuti

Consente di fronteggiare problematiche legate a valori mancanti e outlier:

- Gestisce le interazioni tra le feature
- Può gestire sia dati numerici sia dati categorici
- Lavora bene su grandi quantità di dati
- Efficiente nei tempi di esecuzione

Gli **svantaggi** principali sono i seguenti:

- Maggior tendenza all'overfitting, specialmente quando l'albero di decisione è particolarmente profondo. Questo effetto è dovuto al grado di specificità raggiunto con la sequenza di nodi di decisione, che potrebbe portare a sfruttare un ristretto sottoinsieme di esempi per raggiungere la risposta finale. Un modo per poter fronteggiare questo problema è l'impostazione di un limite massimo di profondità, che da un lato contribuisce a limitare il rischio di overfitting ma, dall'altro lato, aumenta l'errore dovuto ai bias, cioè alle assunzioni effettuate. In questo modo si ottiene un modello più semplice con meno varianza a seconda dei casi considerati, tuttavia un modello predittivo più debole. Idealmente, *si è intenzionati a ridurre sia l'errore dovuto alla presenza di bias, sia l'errore dovuto alla varianza* (tradeoff Bias-Variance).
- Non supportano l'apprendimento online, per cui è necessario ricostruire l'albero di decisione per integrare nuovi esempi
- Richiede molti più dati rispetto a quanti ne richiede il Naive Bayes
- La costruzione di alberi di decisione richiede l'utilizzo di algoritmi in grado di individuare la scelta ottimale tra ciascun nodo. Uno degli algoritmi più utilizzati è quello di *Hunt*. Si tratta di un modello greedy, nel senso che effettua la scelta più ottima in ciascuno step, ma non tiene conto dell'ottimo globale. Ciò significa che, pur tenendo conto dell'alternativa ottima in ciascuno step (ottimo locale), non è detto che il percorso seguito sia quello conducente alla decisione ottima a livello globale

## 2.4 Random forest

Le Random Forests rappresentano un metodo di **ensemble learning** che consente di effettuare compiti di varia tipologia, come la classificazione oppure la regressione. In particolare, il metodo consiste nell'utilizzo di molteplici alberi di decisione durante la fase di apprendimento e, in task di classificazione, restituisce la classe che rappresenta la moda tra le classi dei singoli alberi di decisione.

I **vantaggi** principali sono i seguenti:

- Riducono la tendenza all'overfitting che si verifica nell'utilizzo dei Decision Trees, senza un incremento sostanziale dell'errore dovuto al bias. Questo è possibile apprendendo da differenti dataset di casi oppure utilizzando insiemi di features casuali. In quest'ultimo caso, dato che la Random Forest è un insieme di Decision Trees, se ciascun caso di esempio si forma di 20 features, ciascun *Decision Tree* utilizzerà un gruppo casuale di features (ad esempio, 5) tra le 20 features presenti nel dataset. In questo modo, la Random Forest è in grado di ridurre l'errore dovuto alla presenza di *bias* e l'errore dovuto alla varianza. Risulta fondamentale che le *feature* di ciascun *Decision Tree* siano scelte in maniera casuale per evitare correlazione tra gli alberi di decisione. In caso contrario, le features con una forte influenza sulla predizione sarebbero utilizzate da molteplici alberi di decisione, andando incontro all'errore dovuto alla varianza
- Lavora bene su grandi quantità di dati

Gli **svantaggi** principali sono i seguenti.

- Rispetto al singolo Decision Tree, è più complesso da interpretare e rappresentare
- Richiede un'adeguata parametrizzazione per raggiungere i migliori risultati possibili in termini di accuratezza. In particolare, sono fondamentali i parametri di randomizzazione (selezione dei nodi, numero degli alberi e randomizzazione degli attributi da considerare in ciascun albero)

### 3 Dataset utilizzati

#### 3.1 Scelta dei dataset

La scelta dei dataset più appropriati alla sperimentazione è stata presa basandosi su diversi criteri. Innanzitutto, come spiegato nel *Capitolo 1* è stato fondamentale selezionare dataset provenienti da **contesti applicativi differenti**. Inoltre, hanno influito nella scelta anche il numero di **istanze**, il numero di **attributi**, il numero di **classi** e la presenza/assenza di **dati mancanti**. Le performance di alcuni algoritmi, infatti, potrebbero avere performance diverse fra dataset di piccole dimensioni e su dataset di grandi dimensioni, così come potrebbero emergere comportamenti differenti a seconda che il task di classificazione sia **binario** oppure **multi-classe**.

#### 3.2 Adult dataset

Adult è un dataset realizzato con l'obiettivo di classificare i soggetti in base al reddito annuo, distinguendo quelli con un reddito minore o uguale a 50.000\$ da quelli con un reddito superiore, sfruttando i dati di un censimento. Nel dataset sono presenti 32561 istanze, di cui 24720 appartengono alla classe “**<=50**” e 7841 a “**>50**”.

Come si può dedurre dai dati appena riportati, il dataset **non è bilanciato**, è presente un forte sbilanciamento verso le istanze appartenenti alla classe “**<=50K**”.

Gli attributi di cui si compone ciascuna istanza sono i seguenti:

Nome attributo	Tipologia	Descrizione
Age	Numerico	
Workclass	Categorico	tipologia di lavoro (azienda privata, azienda pubblica, imprenditoria, ...)
Fnlwgt	Numerico	peso utilizzato per scopi statistici
Education	Categorico	
Education.num	Numerico	
Marital.status	Categorico	Stato civile
Occupation	Categorico	Tipologia di occupazione (Vendite, Tech-Support, Forze armate, Trasporti, ...)
Relationship	Categorico	Relazione (moglie, figlio unico, ...)
Race	Categorico	
Sex	Categorico	



Capital.gain	Numerico	
Capital.loss	Numerico	
House.per.week	Numerico	
Native.country	Categorico	

Tabella 1 - Attributi delle istanze di Adult

### 3.3 Vehicle Silhouettes dataset

Il dataset contiene informazioni per discriminare le silhouette di diversi veicoli tra automobili, van e bus ed è composto da 846 istanze.

Nel dataset si possono distinguere 4 classi:

- **Opel:** 212
- **Saab:** 217
- **Bus:** 218
- **Van:** 199

Come si può dedurre dai dati sopra riportati, il dataset è molto **bilanciato**.

Gli attributi di cui si compone ciascuna istanza sono i seguenti:

Nome attributo	Tipologia	Descrizione
Compactness	Numerico	$(\text{average perim})^2 / \text{area}$
Circularity	Numerico	$(\text{average radius})^2 / \text{area}$
Distance circularity	Numerico	$\text{Area} / (\text{av.distance from border})^2$
Radius ratio	Numerico	$(\text{max.rad} - \text{min.rad}) / \text{av.radius}$
Pr.axis aspet ratio	Numerico	$(\text{minor axis}) / (\text{major axis})$
Max length aspect ratio	Numerico	$(\text{length perp. max length}) / (\text{max length})$
Scatter ratio	Numerico	$(\text{inertia about minor axis}) / (\text{inertia about major axis})$
Elongatedness	Numerico	$\text{Area} / (\text{shrink width})^2$
Pr.axis rectangularity	Numerico	$\text{Area} / (\text{pr.axis length} * \text{pr.axis width})$

Max. length rectangularity	Numerico	Area / (max.length*length perp. to this)
Scaled variance_major	Numerico	(2nd order moment about minor axis)/area
Scaled variance_minor	Numerico	(2nd order moment about major axis)/area ALONG MINOR AXIS
Scaled radius of gyration	Numerico	(mavar+mivar) / area
Skewness about major axis	Numerico	(3rd order moment about major axis)/sigma_min**3
Skewness about minor axis	Numerico	(3rd order moment about minor axis) / sigma_maj**3
Kurtosis about minor axis	Numerico	(4th order moment about major axis) / sigma_min**4
Kurtosis about minor axis	Numerico	(4th order moment about minor axis)/sigma_maj**4
Hollows ratio	Numerico	(area of hollows) / (area of bounding polygon)

*Tabella 2 - Attributi delle istanze di Silhouettes*

## 4 Sperimentazione

### 4.1 Fase preliminare

I dataset selezionati per l'esecuzione delle sperimentazioni sono stati reperiti dall'UCI Machine Learning Repository in formato **CSV**, per utilizzarli in WEKA si è ritenuto opportuno effettuare una conversione dei dataset nel formato **ARFF**. Per la conversione è stato utilizzato il tool **ArffViewer**, che ha permesso di assegnare un nome a ciascun attributo e selezionare l'attributo contenente la classe.

### 4.2 Sperimentazione con parametri di default

#### 4.2.1 Adult dataset

Di seguito vengono mostrati e discussi i risultati ottenuti dalla sperimentazione sul dataset "Adult". La validazione del modello è stata effettuata mediante **cross-validation** con **10 fold** e l'esperimento è stato ripetuto 10 volte.

Prima di effettuare il confronto, ciascun algoritmo è stato eseguito singolarmente in **Weka Explorer**

con lo scopo di catturare ulteriori dettagli sui risultati. Successivamente, è stato eseguito il confronto utilizzando **Weka Experiment**, che ha restituito i seguenti risultati in termini di istanze classificate correttamente:

- **Naive Bayes:** 83.42%
- **Decision Tree:** 84.73%
- **Random Forest:** 85.01%

## Naive Bayes

Naive Bayes – Risultati generali	
Tempo per la costruzione del modello	0.04s
Istanze classificate correttamente	27165 (83.43%)
Istanze classificate non correttamente	5395 (16.57%)

*Tabella 3 - Risultati generali del Naive Bayes su Adult*

Naive Bayes – Risultati per classe		
Classe	Precision	Recall
<=50K	0.859	0.935
>50K	0.716	0.518

*Tabella 4 - Risultati per classe del Naive Bayes su Adult*

Naive Bayes – Matrice di confusione		
	<=50K	>50K
<=50K	23105	1614
>50K	3781	4060

*Tabella 5 - Matrice di confusione del Naive Bayes su Adult*

Nonostante la semplicità del modello, i risultati nelle tabelle precedenti mostrano che più dell'83% di istanze sono state classificate correttamente. Tuttavia, la matrice di confusione mette in evidenza le difficoltà riscontrate nella classificazione delle istanze di classe ">50K", con 3781 istanze erroneamente classificate come "<=50K". L'origine di questa problematica è da imputare al forte sbilanciamento del dataset, con 24719 istanze di classe "<=50K" e 7841 istanze di classe ">50K".

Questo sbilanciamento ha avuto ripercussioni sulla precisione di classificazione delle istanze appartenenti alla classe “>50K” (0.716). Inoltre, le assunzioni di cui si caratterizza il Naive Bayes possono aver avuto un’influenza determinante nei risultati ottenuti, causando un fenomeno di *underfitting*.

## Decision Tree

Decision Tree – Risultati generali	
Tempo per la costruzione del modello	4s
Istanze classificate correttamente	27601 (84.77%)
Istanze classificate non correttamente	4959 (15.23%)

Tabella 6 - Risultati generali del Decision Tree su Adult

Decision Tree – Risultati per classe		
Classe	Precision	Recall
<=50K	0.880	0.927
>50K	0.723	0.596

Tabella 7 - Risultati per classe del Decision Tree su Adult

Decision Tree – Matrice di confusione		
	<=50K	>50K
<=50K	22924	1795
>50K	3164	4677

Tabella 8 - Matrice di confusione del Decision Tree su Adult

Osservando le tabelle si può constatare che più dell’85% di istanze viene classificate correttamente. Anche in questo caso, il modello ha avuto difficoltà nella classificazione delle istanze appartenenti alla classe “>50K”. Infatti, è riuscito a raggiungere una precisione di oltre l’87% con le istanze di classe “<=50K”, mentre la precisione di classificazione delle istanze di classe “>50K” si è fermata a circa il 72%. Rispetto ai risultati ottenuti con il Naive Bayes, si nota un leggero miglioramento, che potrebbe essere giustificato dalla capacità dei *Decision Tree* di gestire le interazioni tra le feature.

## Random Forest

Random Forest – Risultati generali	
Tempo per la costruzione del modello	39.04s
Istanze classificate correttamente	27676 (85.00%)
Istanze classificate non correttamente	4884 (15.00%)

Tabella 9 - Risultati generali del Random Forest su Adult

Random Forest – Risultati per classe		
Classe	Precision	Recall
<=50K	0.880	0.930
>50K	0.730	0.598

Tabella 10 - Risultati per classe del Random Forest su Adult

Random Forest – Matrice di confusione		
	<=50K	>50K
<=50K	22984	1735
>50K	3149	4692

Tabella 11 - Matrice di confusione del Random Forest su Adult

I risultati nelle tabelle mostrano ancora una volta dei risultati discreti, infatti vi sono l'85% di istanze classificate correttamente. Anche quest'ultimo modello ha subito gli effetti dello sbilanciamento del dataset, riscontrando difficoltà nella classificazione delle istanze appartenenti alla classe ">50K". I risultati ottenuti sono leggermente inferiori rispetto a quelli ottenuti con il *Decision Tree*, nonostante il *Random Forest* dovrebbe rappresentare un'alternativa valida per rimediare al problema di overfitting presente nel *Decision Tree*. Ovviamente, il *Random Forest* per sua natura richiede più tempo per generare il modello rispetto al *Decision Tree*.

### 4.2.3. Adult Dataset (con preprocessing)

Di seguito vengono mostrati i risultati ottenuti dalle sperimentazioni, successivamente alle operazioni di preprocessing attuate con l'obiettivo di adeguare il dataset agli algoritmi di apprendimento automatico in questione. Sono stati valutati una serie di interventi volti a preparare al meglio il dataset a disposizione:

1. **Sostituzione di valori nulli;**
2. **Rimozione di eventuali attributi superflui;**
3. **Discretizzazione di alcuni attributi continui;**
4. **Bilanciamento del dataset con tecniche di oversampling.**

Per quanto riguarda la sostituzione di **valori nulli**, è stato sfruttato un filtro messo a disposizione da Weka Explorer (*ReplaceMissingValues*) che ha consentito di sostituire i valori nulli con la relativa moda. Il filtro ha agito sui seguenti attributi:

- *workclass* (1836 istanze)
- *occupation* (1843 istanze)
- *native\_country* (857 istanze)

Tuttavia, la sostituzione dei valori nulli non ha aiutato nell'ottenimento di risultati migliori. In particolare, il Naive Bayes non ha ottenuto nessun beneficio in termini di accuratezza dei risultati, mentre il Decision Tree (e di conseguenza anche il Random Forest) è già in grado di gestire le problematiche legate ai valori mancanti. In merito alla rimozione di eventuali **attributi superflui**, è stata effettuata un'analisi per valutare l'eliminazione dell'attributo *fnlwgt*, un peso calcolato sulla base delle caratteristiche demografiche del soggetto. Tuttavia, anche in questo caso l'operazione non ha restituito benefici nei risultati di classificazione.

Si è prestata particolare cura all'operazione di **discretizzazione**, infatti ogni operazione di questo genere comporta una perdita di informazioni, col rischio di compromettere i risultati di classificazione. La discretizzazione è stata testata su tutti gli attributi continui delle istanze, tuttavia non sono emersi particolari vantaggi dal punto di vista della classificazione. Nelle seguenti sono mostrati i risultati ottenuti, prima con la discretizzazione dell'attributo *age* e dopo con la discretizzazione dell'attributo *hours.per.week*. Utilizzando il Naive Bayes sono emersi miglioramenti di lieve entità, lo stesso non si è riscontrato con l'utilizzo del *Decision Tree*. La ragione di questo appena affermato risiede nel fatto che i *Decision Tree* necessitano di grandi quantità di dati per restituire risultati di classificazione soddisfacenti e la discretizzazione ha sempre come risolto la perdita di informazioni.

Tuttavia, la discretizzazione può tornare utile nel rendere il modello più semplice ed interpretabile, soprattutto nel caso di alberi di decisione complessi. Il Naive Bayes ha giovato della discretizzazione, poiché seppur in presenza di pochi dati è in grado di raggiungere risultati soddisfacenti.

Naive Bayes – Risultati generali	
Tempo per la costruzione del modello	0.06s
Istanze classificate correttamente	27179 (83.47%)
Istanze classificate non correttamente	5381 (16.53%)

Tabella 12 - Risultati generali del NB su Adult (discretizzazione di "age")

Naive Bayes – Risultati per classe		
Classe	Precision	Recall
<=50K	0.860	0.934
>50K	0.715	0.521

Tabella 13 - Risultati per classe del NB su Adult (discretizzazione di "age")

Naive Bayes – Matrice di confusione		
	<=50K	>50K
<=50K	23090	1629
>50K	3752	4089

Tabella 14 - Matrice di confusione del NB su Adult (discretizzazione di "age")

Naive Bayes – Risultati generali	
Tempo per la costruzione del modello	0.06s
Istanze classificate correttamente	27209 (83.57%)
Istanze classificate non correttamente	5351 (16.43%)

Tabella 15 - Risultati generali del NB su Adult (discretizzazione di "hours.per.week")

Naive Bayes – Risultati per classe		
Classe	Precision	Recall
<=50K	0.862	0.933
>50K	0.715	0.529

Tabella 16 - Risultati per classe del NB su Adult (discretizzazione di "hours.per.weeks")

Decision Tree – Risultati generali	
Tempo per la costruzione del modello	0.64s
Istanze classificate correttamente	27492 (84.43%)
Istanze classificate non correttamente	5068 (15.57%)

Tabella 17 - Risultati generali del DT su Adult (discretizzazione di "hours.per.week")

Decision Tree – Risultati per classe		
Classe	Precision	Recall
<=50K	0.876	0.926

>50K	0.715	0.588
------	-------	-------

Tabella 18 - Risultati per classe del DT su Adult (discretizzazione di "hours.per.week")

Decision Tree – Matrice di confusione		
	<=50K	>50K
<=50K	22881	1838
>50K	3230	4611

Tabella 19 – Matrice di confusione del DT su Adult (discretizzazione di "hours.per.week")

I risultati emersi applicando la discretizzazione di hours\_per\_week all'albero decisionale non saranno riportati, in quanto totalmente identici all'applicazione della discretizzazione su age.

Con lo scopo di rendere più bilanciato il dataset, è stata utilizzata la tecnica **SMOTE** (*Synthetic Minority Over-sampling Technique*), che implementa un approccio per la costruzione di classificatori a partire da dataset di training non bilanciati, ovvero dataset in cui le classi non sono equamente rappresentate. Nello specifico, data un'istanza della classe di cui si vuole effettuare l'*oversampling*, per fare ciò, vengono considerate le k istanze più vicine da combinare per poter generare nuove istanze. SMOTE è stato utilizzato con i parametri di default (k=5), impostando 200% come percentuale di istanze da creare. Il risultato è un dataset bilanciato di 48242 istanze, di cui 24719 di classe "<=50K" e 23523 di classe ">50K". Il bilanciamento ha portato a significativi miglioramenti nei risultati della classificazione



Naive Bayes – Risultati generali	
Tempo per la costruzione del modello	0.04s
Istanze classificate correttamente	40997 (84.98%)
Istanze classificate non correttamente	7245 (15.02%)

Tabella 20 - Risultati generali del NB su Adult (dopo oversampling)

Naive Bayes – Risultati per classe		
Classe	Precision	Recall
<=50K	0.813	0.918
>50K	0.900	0.779

Tabella 21 - Risultati per classe del NB su Adult (dopo oversampling)

Naive Bayes – Matrice di confusione		
	<=50K	>50K
<=50K	22681	2038
>50K	5207	18316

Tabella 22 – Matrice di confusione del NB su Adult (dopo oversampling)

Decision Tree – Risultati generali	
Tempo per la costruzione del modello	2.7s
Dimensione dell'albero	1872
Istanze classificate correttamente	42806 (88.73%)
Istanze classificate non correttamente	5436 (11.27%)

Tabella 23 - Risultati generali del DT su Adult (dopo oversampling)

Decision Tree – Risultati per classe		
Classe	Precision	Recall
<=50K	0.878	0.906
>50K	0.898	0.867

Tabella 24 - Risultati per classe del DT su Adult (dopo oversampling)

Decision Tree – Matrice di confusione		
	<=50K	>50K
<=50K	22403	2316

>50K	3120	20403
------	------	-------

Tabella 25 – Matrice di confusione del DT su Adult (dopo oversampling)

Random Forest – Risultati generali	
Tempo per la costruzione del modello	23.55s
Istanze classificate correttamente	43196 (89.54%)
Istanze classificate non correttamente	5046 (10.46%)

Tabella 26 - Risultati generali del RF su Adult (dopo oversampling)

Random Forest – Risultati per classe		
Classe	Precision	Recall
<=50K	0.883	0.918
>50K	0.910	0.872

Tabella 27 - Risultati per classe del RF su Adult (dopo oversampling)

Random Forest – Matrice di confusione		
	<=50K	>50K
<=50K	22694	2025
>50K	3021	20502

Tabella 28 – Matrice di confusione del RF su Adult (dopo oversampling)

#### 4.2.2 Silhouettes dataset

Di seguito vengono mostrati e discussi i risultati ottenuti dalla sperimentazione sul dataset “Silhouettes”. La validazione del modello è stata effettuata mediante **cross-validation** con **10 fold** e l’esperimento è stato ripetuto 10 volte.

Prima di effettuare il confronto, ciascun algoritmo è stato eseguito singolarmente in **Weka Explorer** con lo scopo di catturare ulteriori dettagli sui risultati. Successivamente, è stato eseguito il confronto utilizzando **Weka Experiment**, che ha restituito i seguenti risultati in termini di istanze classificate correttamente:

- **Naive Bayes:** 44.68%
- **Decision Tree:** 70.18%
- **Random Forest:** 74.87%

## Naive Bayes

Naive Bayes – Risultati generali	
Tempo per la costruzione del modello	0.02s
Istanze classificate correttamente	379 (44.79%)
Istanze classificate non correttamente	420 (55.20%)

Tabella 26 - Risultati generali del RF su Silhouettes

Naive Bayes – Risultati per classe		
Classe	Precision	Recall
Opel	0.447	0.415
Saab	0.531	0.392
Bus	0.653	0.147
Van	0.395	0.874

Tabella 27 - Risultati per classe del RF su Silhouettes

Naive Bayes – Matrice di confusione				
	Opel	Saab	Bus	Van
Opel	88	60	0	64
Saab	62	85	3	67
Bus	43	8	32	135
Van	4	7	14	174

Tabella 28 – Matrice di confusione per classe del RF su Silhouettes

Dai dati riportati emergono chiare problematiche legate all'utilizzo del Naive Bayes su questo dataset, ciò si deduce dalla percentuale di classificazioni errate e viene ulteriormente evidenziato dalla matrice di confusione dalla quale si evincono difficoltà soprattutto nel distinguere tra le istanze della classe "opel" e quelle della classe "saab".

## Decision tree

Decision tree – Risultati generali	
Tempo per la costruzione del modello	0.07s
Istanze classificate correttamente	613 (72.45%)
Istanze classificate non correttamente	233 (27.54%)

Tabella 29 - Risultati generali del DT su Silhouettes

Decision tree – Risultati per classe		
Classe	Precision	Recall
Opel	0.553	0.613
Saab	0.524	0.456
Bus	0.945	0.950
Van	0.872	0.889

Tabella 30 - Risultati per classe del DT su Silhouettes

Decision tree – Matrice di confusione				
	Opel	Saab	Bus	Van
Opel	130	74	2	6
Saab	96	99	6	16
Bus	3	4	207	4
Van	6	12	4	177

Tabella 31 – Matrice di confusione per classe del DT su Silhouettes

Nel Decision Tree le prestazioni risultano significativamente incrementate rispetto a quelle ottenute con l'utilizzo del Naive Bayes. Tuttavia, dalla matrice di confusione, si evince ancora una volta come vi siano problematiche relativi alla distinzione tra entità della classe “opel” e della classe “saab”.

## Random Forest

Random Forest – Risultati generali	
Tempo per la costruzione del modello	0.26s
Istanze classificate correttamente	643 (76.00%)
Istanze classificate non correttamente	203 (23.99%)

Tabella 32 - Risultati generali del RF su Silhouettes

Random Forest – Risultati per classe		
Classe	Precision	Recall
Opel	0.598	0.547
Saab	0.573	0.562
Bus	0.960	0.982
Van	0.884	0.960

Tabella 33 - Risultati per classe del RF su Silhouettes

	Random Forest – Matrice di confusione			
	Opel	Saab	Bus	Van
Opel	116	84	2	10
Saab	76	122	6	13
Bus	0	2	214	2
Van	2	5	1	191

Tabella 34 – Matrice di confusione per classe del RF su Silhouettes

### 4.2.3. Silhouettes Dataset (con preprocessing)

Di seguito vengono mostrati i risultati ottenuti dalle sperimentazioni, successivamente alle operazioni di preprocessing attuate con l'obiettivo di adeguare il dataset agli algoritmi di apprendimento automatico in questione. A differenza del dataset Adult, in questo dataset non è stato necessario effettuare l'eliminazione dei **valori nulli** in quanto già assenti. Il dataset Silhouettes inoltre risulta essere già bilanciato, quindi non è stato necessario neanche effettuare operazioni col fine di bilanciare i dati. Tuttavia, il dataset presenta diversi attributi con valori continui molto sparsi, quindi si è ritenuto opportuno effettuare una discretizzazione. Nello specifico, gli attributi discretizzati sono stati:

- pr.axis aspect ratio

- max length\_aspect\_ratio
- skewness\_about\_major

Il processo di discretizzazione ha portato significativi miglioramenti solo ai risultati del Naive Bayes, i quali tuttavia rimangono decisamente non soddisfacenti. Nel caso del Decision Tree e del Random Forest non sono stati riscontrati cambiamenti significativi.

Naive Bayes – Risultati generali	
Tempo per la costruzione del modello	0.02s
Istanze classificate correttamente	530 (62.64%)
Istanze classificate non correttamente	316 (37.35%)

Tabella 35 - Risultati generali del NB su Silhouettes dopo la discretizzazione

Naive Bayes – Risultati per classe		
Classe	Precision	Recall
Opel	0.510	0.389
Saab	0.520	0.484
Bus	0.822	0.720
Van	0.630	0.975

Tabella 36 - Risultati per classe del NB su Silhouettes dopo la discretizzazione

Naive Bayes – Matrice di confusione				
	Opel	Saab	Bus	Van
Opel	74	74	13	51
Saab	42	105	16	54
Bus	29	23	157	9
Van	0	0	5	194

Tabella 37 – Matrice di confusione per classe del NB su Silhouettes dopo la discretizzazione

## 4.3 Sperimentazione con parametrizzazione

### 4.3.1 Tuning dei parametri

In questa sezione vengono mostrati i risultati raggiunti successivamente ad un tuning dei parametri degli algoritmi di apprendimento considerati. Per un'adeguata parametrizzazione sono state considerate due vie alternative:

- utilizzo di un package Weka, **Auto-WEKA**, che permette, a partire da un dataset di riferimento, di individuare il modello Weka più adeguato, con relativi parametri;
- utilizzo di un altro meta-classificatore, **CVParameterSelection**, che permette, a partire da un dataset di riferimento e da un algoritmo di classificazione, di individuare i parametri più adeguati.

Il meta-classificatore *CVParameterSelection* è stato ritenuto più utile per gli scopi specifici in quanto si vuole individuare i parametri più adeguati per il *Decision Tree* e il *Random Forest*. La parametrizzazione è stata effettuata considerando il dataset successivo alle operazioni di preprocessing. Il test è stato effettuato con **10-fold cross validation** per il *Decision Tree* e con **5-fold cross validation** per il *Random Forest* (per motivi di efficienza).

### 4.3.2. Adult Dataset

Il tuning dei parametri ha permesso di migliorare ulteriormente i risultati di classificazione, già considerati più che soddisfacenti. Sono state compiute molteplici ricerche, utilizzando diversi parametri. Per quanto riguarda il **Decision Tree**, i parametri sui quali il tuning ha fornito risvolti positivi sono:

- **M**: minimo numero di istanze per foglia (default: 2)
- **L**: profondità massima dell'albero (default, nessun vincolo)

Il *CVParameterSelection*, con il Decision Tree, è stato eseguito impostando i seguenti intervalli di valori di test:

- **M**: 1 10, con 10 step
- **L**: 1 20, con 20 step

Il tuning dei parametri applicato al Decision Tree non ha comportato il miglioramento dei risultati di classificazione, ma ha permesso una leggera **semplificazione dell'albero di decisione**. Infatti, l'albero di decisione creato ha dimensione (numero di nodi) pari a 1810 (contro i 1872 precedenti).

Per quanto riguarda il **Random Forest**, i parametri sui quali il tuning ha fornito risvolti positivi sono:

- **M**: minimo numero di istanze per foglia (default: 1)

- **Depth:** massima profondità degli alberi di decisione (default, nessun vincolo)

Il *CVParameterSelection*, con il Random Forest, è stato eseguito impostando i seguenti intervalli di valori di test:

- **M:** 1 10, con 10 step
- **Depth:** 1 10, con 10 step

Non è stato fattibile effettuare il tuning su un terzo parametro, cioè il numero di alberi da costruire (**I**), a causa dei lunghi tempi richiesti. Di default, il Random Forest di Weka effettua 100 iterazioni, quindi costruisce 100 alberi decisionali. Anche nel Random Forest non ci sono stati particolari miglioramenti nei risultati di classificazione

Nelle seguenti tabelle sono mostrati i risultati ottenuti con il Decision Tree e con il Random Forest.

Decision tree – Parametri e risultati generali	
<b>M (minimo numero di istanze per foglia)</b>	2
<b>L (profondità massima dell'albero)</b>	15
<b>Dimensione dell'albero</b>	1810
<b>Tempo per la costruzione del modello</b>	27,2s
<b>Istanze classificate correttamente</b>	42755 (88,63%)
<b>Istanze classificate non correttamente</b>	5487 (11,37%)

*Tabella 38 - Matrice di confusione del RF su Adult (con parametrizzazione)*

Decision Tree – Risultati per classe		
Classe	Precision	Recall
<b>&lt;=50K</b>	0.879	0.902
<b>&gt;50K</b>	0.894	0.870

*Tabella 39 - Risultati per classe del DT su Adult (con parametrizzazione)*

Decision Tree – Matrice di confusione		
	<b>&lt;=50K</b>	<b>&gt;50K</b>
<b>&lt;=50K</b>	22298	2421
<b>&gt;50K</b>	3066	20457

*Tabella 40 – Matrice di confusione del DT su Adult (con parametrizzazione)*



Random Forest – Parametri e risultati generali	
M (minimo numero di istanze per foglia)	7
Depth (profondità massima dell'albero)	10
Tempo per la costruzione del modello	19,1s
Istanze classificate correttamente	43327(89,81%)
Istanze classificate non correttamente	4915(10,31%)

Tabella 41 - Matrice di confusione del RF su Adult (con parametrizzazione)

Random Forest – Risultati per classe		
Classe	Precision	Recall
<=50K	0.894	0.908
>50K	0.902	0.887

Tabella 42 - Risultati per classe del RF su Adult (con parametrizzazione)

Random Forest – Matrice di confusione		
	<=50K	>50K
<=50K	22453	2266
>50K	2649	20874

Tabella 43 – Matrice di confusione del RF su Adult (con parametrizzazione)

#### 4.3.2 Silhouettes dataset

Anche per il dataset Silhouette si è tentato di effettuare una fase di tuning utilizzando **CVParameterSelection**, ma non sono stati individuati parametri, per cui esistessero valori in grado di migliorare le prestazioni.

#### 4.4. Test di valutazione

I modelli vengono comunemente valutati utilizzando metodi di campionamento come la k-fold cross-validation, in cui vengono calcolati e confrontati i punteggi medi. Tuttavia, questo approccio può essere fuorviante in quanto i punteggi medi potrebbero essere reali oppure il risultato di un colpo di fortuna statistico. Per affermare quale dei classificatori considerati risulti essere il migliore per i problemi trattati è necessario effettuare dei test di significatività statistica, come il **Paired t-test**. Il Paired t-test è una procedura statistica per comparare i risultati di misurazione ottenuti su un gruppo

utilizzando 2 classificatori differenti. Considerando media e varianza delle differenze ottenute nelle 2 misurazioni nel corso di n esecuzioni, viene calcolato il t-value. Si ottiene la **probabilità che le 2 misurazioni siano significativamente differenti** utilizzando il t-value ed un livello di importanza (*significance*), normalmente impostato al 5%.

La valutazione è stata effettuata sui dataset risultanti dalle operazioni di preprocessing, con l'utilizzo dei parametri individuati durante la fase di ottimizzazione. Ricapitolando, i dataset coinvolti sono:

- **Adult**
  - o Numero istanze: 72216
  - Classe “<=50K”: 37155
  - Classe “>50K”: 35061 (con oversampling)
  - o Parametri ottimali individuati (Decision Tree)
    - M (Minimo numero di istanze per foglia): 2
    - L (Profondità massima dell'albero): 15
  - o Parametri ottimali individuati (Random Forest)
    - M (Minimo numero di istanze per foglia): 7
    - Depth (Profondità massima degli alberi): 10
- **Silhouettes**
  - o Numero istanze: 846
  - Classe “opel”: 212
  - Classe “saab”: 217
  - Classe “bus”: 218
  - Classe “van”: 199

I parametri ideali individuati sono per questo dataset sono quelli di default.

L'esperimento, i cui risultati sono mostrati di seguito, è stato condotto per ciascun dataset utilizzando la **cross validation con 10 fold** (*experiment type*) ed effettuando **10 ripetizioni** (*iteration control*). L'analisi dei risultati con **Paired T-Tester** è stata condotta con una *significance* del 5%.

I dati mostrati nella seguente tabella rappresentano la percentuale di istanze classificate correttamente da ciascun classificatore, per ciascun dataset considerato.

<b>Adult Dataset</b>	
Naive Bayes	84.37%
Decision Tree	88.58%
Random Forest	<b>89.85%</b>
<b>Silhouettes Dataset</b>	
Naive Bayes	62.63%
Decision Tree	70.18%
Random Forest	<b>74.87%</b>

*Tabella 44 - Percentuale di istanze classificate correttamente per ciascun dataset e classificatore*

La seguente tabella contiene il risultato del confronto tra i classificatori utilizzati.

	<b>Naive Bayes</b>	<b>Decision Tree</b>	<b>Random Forest</b>
<b>Naive Bayes</b>	-	2	2
<b>Decision Tree</b>	0	-	2
<b>Random Forest</b>	0	0	-

*Tabella 45 - Confronto tra classificatori*

Ciascuna delle celle contiene il numero dei dataset sui quali il classificatore posto nella colonna ha funzionato meglio del classificatore posto nella riga (secondo il Paired t-test). Con i dataset considerati, il Decision Tree si dimostra statisticamente migliore rispetto al Naive Bayes e il Random Forest si dimostra statisticamente migliore rispetto al Decision Tree. Questa conclusione è in parte giustificata dalle peculiarità di ciascun classificatore:

- il Naive Bayes assume che gli attributi siano indipendenti tra loro, aumentando il rischio di errori di classificazione, mentre il Decision Tree è in grado di gestire le interazioni tra gli attributi
- il Decision Tree tende maggiormente all'*overfitting* a seconda del grado di specificità dell'albero di decisione, mentre il Random Forest riduce la tendenza all'*overfitting*

utilizzando molteplici Decision Tree, ciascuno dei quali è generato da un dataset differente oppure utilizzando gruppi di attributi casuali

## 5 Modelli generati

### 5.1 Modello di Adult con il Naive Bayes

Attribute	Class	
	<=50K	>50K
	(0.51)	(0.49)
age		
mean	36.7836	44.1956
std. dev.	14.0201	9.2663
weight sum	24719	23523
precision	0.0048	0.0048
workclass		
Self-emp-not-inc	1818.0	1731.0
Private	17734.0	19256.0
State-gov	945.0	404.0
Federal-gov	590.0	450.0
Local-gov	1477.0	734.0
Self-emp-inc	495.0	763.0
Without-pay	15.0	1.0
Never-worked	8.0	1.0
[total]	23082.0	23340.0
fnlwgt		
mean	190345.5056	187952.6466
std. dev.	106480.0288	102198.3359
weight sum	24719	23523
precision	40.7816	40.7816
education		
Bachelors	3134.0	10390.0
HS-grad	8827.0	5592.0
11th	1116.0	65.0
Masters	765.0	2082.0
9th	488.0	28.0

Some-college	5905.0	3316.0
Assoc-acdm	803.0	306.0
Assoc-voc	1022.0	476.0
7th-8th	607.0	41.0
Doctorate	108.0	423.0
Prof-school	154.0	694.0
5th-6th	318.0	17.0
10th	872.0	65.0
1st-4th	163.0	7.0
Preschool	52.0	1.0
12th	401.0	36.0
[total]	24735.0	23539.0

education.num

mean	9.5948	11.6151
std. dev.	2.436	2.1114
weight sum	24719	23523
precision	0.0012	0.0012

marital.status

Married-civ-spouse	8285.0	22281.0
Divorced	3981.0	486.0
Married-spouse-absent	385.0	39.0
Never-married	10192.0	560.0
Separated	960.0	67.0
Married-AF-spouse	14.0	11.0
Widowed	909.0	86.0
[total]	24726.0	23530.0

occupation

Exec-managerial	2099.0	10031.0
Handlers-cleaners	1285.0	105.0
Prof-specialty	2282.0	6286.0
Other-service	3159.0	168.0
Adm-clerical	3263.0	826.0
Sales	2668.0	2296.0
Craft-repair	3171.0	2044.0
Transport-moving	1278.0	427.0

Farming-fishing	880.0	142.0
Machine-op-inspct	1753.0	353.0
Tech-support	646.0	404.0
Protective-serv	439.0	260.0
Armed-Forces	9.0	2.0
Priv-house-serv	149.0	2.0
[total]	23081.0	23346.0

#### relationship

Husband	7276.0	21199.0
Not-in-family	7449.0	1121.0
Wife	824.0	884.0
Own-child	5002.0	68.0
Unmarried	3229.0	219.0
Other-relative	945.0	38.0
[total]	24725.0	23529.0

#### race

White	20699.0	22766.0
Black	2738.0	396.0
Asian-Pac-Islander	764.0	303.0
Amer-Indian-Eskimo	276.0	37.0
Other	247.0	26.0
[total]	24724.0	23528.0

#### sex

Male	15128.0	22169.0
Female	9593.0	1356.0
[total]	24721.0	23525.0

#### capital.gain

mean	148.8552	3976.765
std. dev.	963.9679	14546.8436
weight sum	24719	23523
precision	52.8815	52.8815

#### capital.loss

mean	53.1461	189.857
------	---------	---------

std. dev.	310.7642	583.5323
weight sum	24719	23523
precision	3.0982	3.0982
house.per.week		
mean	38.8385	45.4274
std. dev.	12.3188	9.6244
weight sum	24719	23523
precision	0.008	0.008
native.country		
United-States	21999.0	22854.0
Cuba	71.0	26.0
Jamaica	72.0	11.0
India	61.0	41.0
Mexico	611.0	34.0
South	65.0	17.0
Puerto-Rico	103.0	13.0
Honduras	13.0	2.0
England	61.0	31.0
Canada	83.0	40.0
Germany	94.0	45.0
Iran	26.0	19.0
Philippines	138.0	62.0
Italy	49.0	26.0
Poland	49.0	13.0
Columbia	58.0	3.0
Cambodia	13.0	8.0
Thailand	16.0	4.0
Ecuador	25.0	5.0
Laos	17.0	3.0
Taiwan	32.0	21.0
Haiti	41.0	5.0
Portugal	34.0	5.0
Dominican-Republic	69.0	3.0
El-Salvador	98.0	10.0
France	18.0	13.0

Guatemala	62.0	4.0
China	56.0	21.0
Japan	39.0	25.0
Yugoslavia	11.0	7.0
Peru	30.0	3.0
Outlying-US(Guam-USVI-etc)	15.0	1.0
Scotland	10.0	4.0
Trinidad&Tobago	18.0	3.0
Greece	22.0	9.0
Nicaragua	33.0	3.0
Vietnam	63.0	6.0
Hong	15.0	7.0
Ireland	20.0	6.0
Hungary	11.0	4.0
Holand-Netherlands	2.0	1.0
[total]	24323.0	23418.0

## 5.2 Modello di Adult con il Decision Tree

relationship = Husband

```

| occupation = Exec-managerial
| | education = Bachelors : >50K (3475.39/125.95) [1667.3/50.08]
| | education = HS-grad
| | | education.num < 9
| | | | workclass = Self-emp-not-inc
| | | | capital.loss < 742.5
| | | | | education.num < 8.8 : >50K (5/0) [2/0]
| | | | | education.num >= 8.8
| | | | | capital.gain < 7220.12
| | | | | | fnlwgt < 30102 : >50K (4.22/0.11) [1.06/0.06]
| | | | | | fnlwgt >= 30102
| | | | | | | age < 35.5 : <=50K (11.72/0) [6.22/1]
| | | | | | | age >= 35.5
| | | | | | | | house.per.week < 33.5 : <=50K (10.16/0.11) [4.05/0.06]
| | | | | | | | house.per.week >= 33.5
| | | | | | | | fnlwgt < 212091.5
| | | | | | | | fnlwgt < 181984.5

```



```

| | | | | | | | | | | | | | house.per.week < 36.33 : <=50K (4.06/2) [0/0]
| | | | | | | | | | | | | | house.per.week >= 36.33
| | | | | | | | | | | | | | fnlwgt < 86043 : <=50K (8.39/3.17) [3.17/2.06]
| | | | | | | | | | | | | | fnlwgt >= 86043 : <=50K (12.33/0.33) [8.78/5.22]
| | | | | | | | | | | | | | fnlwgt >= 181984.5 : >50K (7.55/1.28) [3.22/1.22]
| | | | | | | | | | | | | | fnlwgt >= 212091.5 : <=50K (10.61/1.06) [3.05/0.17]
| | | | | | | capital.gain >= 7220.12 : >50K (4/0) [1/0]
| | | | | capital.loss >= 742.5 : >50K (7.22/0.11) [4.11/0]
| | | | workclass = Private
| | | | | house.per.week < 16.79 : <=50K (8.68/0.24) [2.66/1]
| | | | | house.per.week >= 16.79 : >50K (355.05/74.32) [199.95/47.35]
| | | | workclass = State-gov : >50K (7.56/2.45) [2.32/2.27]
| | | | workclass = Federal-gov : >50K (7.56/2.45) [1.32/0.27]
| | | | workclass = Local-gov
| | | | | capital.gain < 4498
| | | | | age < 62.5
| | | | | | | fnlwgt < 264278
| | | | | | | | house.per.week < 46.5
| | | | | | | | | fnlwgt < 164448.5
| | | | | | | | | | fnlwgt < 58338.5 : >50K (2.03/0.03) [1.02/0.02]
| | | | | | | | | | fnlwgt >= 58338.5 : <=50K (3.34/1.08) [0.09/0.02]
| | | | | | | | | | fnlwgt >= 164448.5 : >50K (3.31/0.25) [0.29/0.25]
| | | | | | | | | house.per.week >= 46.5 : <=50K (5.15/1.05) [2.03/1.02]
| | | | | | | | fnlwgt >= 264278 : <=50K (2.22/0.02) [0.15/0.02]
| | | | | | | age >= 62.5 : <=50K (3.66/0.11) [2.43/0.08]
| | | | | capital.gain >= 4498 : >50K (4.03/0.02) [0/0]
| | | | workclass = Self-emp-inc : >50K (62.6/22.74) [23.69/14.24]
| | | | workclass = Without-pay : >50K (0/0) [0/0]
| | | | workclass = Never-worked : >50K (0/0) [0/0]
| | | education.num >= 9 : >50K (958/-0) [428/0]
| | | education = 11th
| | | capital.gain < 4914.5 : <=50K (14.65/3) [7.43/0]
| | | capital.gain >= 4914.5 : >50K (4/-0) [2/0]
| | | education = Masters : >50K (607.08/32.22) [335.22/9.43]
| | | education = 9th : <=50K (4.15/0.36) [6.51/1]
| | | education = Some-college
| | | education.num < 10.01
| | | | education.num < 9.99 : >50K (116/0) [70/0]
| | | | education.num >= 9.99
| | | | capital.gain < 4198.66
| | | | | age < 29.5 : <=50K (25.72/5) [13.36/5]

```

```

| | | | | age >= 29.5
| | | | | native.country = United-States
| | | | | race = White
| | | | | workclass = Self-emp-not-inc
| | | | | fnlwgt < 30276.48 : >50K (7/0) [2/0]
| | | | | fnlwgt >= 30276.48
| | | | | fnlwgt < 70731 : <=50K (4.11/0) [0.06/0]
| | | | | fnlwgt >= 70731
| | | | | capital.loss < 943.5
| | | | | age < 61
| | | | | house.per.week < 59.63 : >50K (20.72/4.5) [8.44/6.28]
| | | | | house.per.week >= 59.63 : <=50K (6/2) [1.06/0]
| | | | | age >= 61 : <=50K (6.05/1.28) [2.44/0.06]
| | | | | capital.loss >= 943.5 : >50K (7.06/0.06) [2.06/0]
| | | | | workclass = Private
| | | | | age < 61.5 : >50K (197.19/41.28) [99.05/26.37]
| | | | | age >= 61.5 : <=50K (12.79/4.14) [6.28/1.46]
| | | | | workclass = State-gov : <=50K (11.43/3.11) [1.24/0.06]
| | | | | workclass = Federal-gov : >50K (8.32/1.23) [6.17/0.13]
| | | | | workclass = Local-gov
| | | | | age < 41.5 : <=50K (6.08/2) [3.02/1]
| | | | | age >= 41.5
| | | | | age < 46 : >50K (5/0) [0.05/0.03]
| | | | | age >= 46 : <=50K (4.51/2.15) [1.25/0.07]
| | | | | workclass = Self-emp-inc
| | | | | fnlwgt < 201735
| | | | | house.per.week < 42.5 : <=50K (4.72/0.11) [5.36/2.07]
| | | | | house.per.week >= 42.5 : >50K (14.01/5.04) [8.11/1.07]
| | | | | fnlwgt >= 201735 : >50K (14.5/1.29) [4.22/0.14]
| | | | | workclass = Without-pay : >50K (0/0) [0/0]
| | | | | workclass = Never-worked : >50K (0/0) [0/0]
| | | | | race = Black
| | | | | fnlwgt < 305682.5 : >50K (5.36/2.36) [2.36/0]
| | | | | fnlwgt >= 305682.5 : <=50K (4/0) [2/2]
| | | | | race = Asian-Pac-Islander : <=50K (2.97/0) [2.69/0.36]
| | | | | race = Amer-Indian-Eskimo : <=50K (1/0) [2/0]
| | | | | race = Other : >50K (0/0) [0/0]
| | | | | native.country = Cuba : >50K (3.02/0.01) [0.01/0.01]
| | | | | native.country = Jamaica : <=50K (1.01/0) [1/1]
| | | | | native.country = India : >50K (0/0) [0/0]
| | | | | native.country = Mexico : >50K (0/0) [0/0]

```

| | | | | | native.country = South : >50K (1.37/0) [0/0]  
 | | | | | | native.country = Puerto-Rico : <=50K (1.01/0) [0/0]  
 | | | | | | native.country = Honduras : >50K (0/0) [0/0]  
 | | | | | | native.country = England : >50K (0/0) [0/0]  
 | | | | | | native.country = Canada : <=50K (0.36/0) [0/0]  
 | | | | | | native.country = Germany : >50K (1.01/1) [2/0]  
 | | | | | | native.country = Iran : >50K (0/0) [0/0]  
 | | | | | | native.country = Philippines : >50K (0.36/0) [0/0]  
 | | | | | | native.country = Italy : >50K (0/0) [0/0]  
 | | | | | | native.country = Poland : >50K (0/0) [0/0]  
 | | | | | | native.country = Columbia : <=50K (1.01/0) [0/0]  
 | | | | | | native.country = Cambodia : >50K (0/0) [0/0]  
 | | | | | | native.country = Thailand : >50K (0/0) [1/0]  
 | | | | | | native.country = Ecuador : >50K (0/0) [0/0]  
 | | | | | | native.country = Laos : >50K (0/0) [0/0]  
 | | | | | | native.country = Taiwan : <=50K (0/0) [1/0]  
 | | | | | | native.country = Haiti : >50K (0/0) [0/0]  
 | | | | | | native.country = Portugal : >50K (0/0) [0/0]  
 | | | | | | native.country = Dominican-Republic : >50K (0/0) [0/0]  
 | | | | | | native.country = El-Salvador : >50K (0/0) [0/0]  
 | | | | | | native.country = France : >50K (0/0) [0/0]  
 | | | | | | native.country = Guatemala : >50K (0/0) [0/0]  
 | | | | | | native.country = China : <=50K (1.01/0) [0/0]  
 | | | | | | native.country = Japan : >50K (0/0) [0/0]  
 | | | | | | native.country = Yugoslavia : >50K (0/0) [1/0]  
 | | | | | | native.country = Peru : >50K (0/0) [0/0]  
 | | | | | | native.country = Outlying-US(Guam-USVI-etc) : >50K (0/0) [0/0]  
 | | | | | | native.country = Scotland : >50K (0/0) [0/0]  
 | | | | | | native.country = Trinidad&Tobago : >50K (0/0) [0/0]  
 | | | | | | native.country = Greece : >50K (0/0) [0/0]  
 | | | | | | native.country = Nicaragua : <=50K (0.36/0) [0/0]  
 | | | | | | native.country = Vietnam : >50K (0/0) [0/0]  
 | | | | | | native.country = Hong : >50K (0/0) [0/0]  
 | | | | | | native.country = Ireland : >50K (0/0) [0/0]  
 | | | | | | native.country = Hungary : >50K (0/0) [0/0]  
 | | | | | | native.country = Holand-Netherlands : >50K (0/0) [0/0]  
 | | | | capital.gain >= 4198.66 : >50K (37.07/0) [18.72/0]  
 | | education.num >= 10.01 : >50K (379/-0) [222/0]  
 | | education = Assoc-acdm : >50K (58.15/8.43) [31.43/9.72]  
 | | education = Assoc-voc  
 | | education.num < 11.04

| | | education.num < 11 : >50K (11/0) [5/0]  
 | | | education.num >= 11  
 | | | | age < 35 : <=50K (5.72/0.36) [5/3]  
 | | | | age >= 35  
 | | | | | house.per.week < 41.5 : >50K (22.07/10.36) [13.72/5.72]  
 | | | | | house.per.week >= 41.5  
 | | | | | | workclass = Self-emp-not-inc : <=50K (1.06/0.06) [2/0]  
 | | | | | | workclass = Private : >50K (20.13/1) [4/0]  
 | | | | | | workclass = State-gov : >50K (0/0) [0/0]  
 | | | | | | workclass = Federal-gov : >50K (0/0) [0/0]  
 | | | | | | workclass = Local-gov : >50K (1.06/0) [0/0]  
 | | | | | | workclass = Self-emp-inc : >50K (3.18/1) [2/0]  
 | | | | | | workclass = Without-pay : >50K (0/0) [0/0]  
 | | | | | | workclass = Never-worked : >50K (0/0) [0/0]  
 | | | education.num >= 11.04 : >50K (26/0) [5/0]  
 | | education = 7th-8th : <=50K (19.23/5.36) [6.15/0.36]  
 | | education = Doctorate : >50K (53.86/1.36) [27.43/3.36]  
 | | education = Prof-school  
 | | | age < 60.82 : >50K (39.36/1.36) [31/3]  
 | | | age >= 60.82  
 | | | | capital.gain < 3209 : <=50K (4.79/0.72) [2.43/0.72]  
 | | | | capital.gain >= 3209 : >50K (4.72/0) [2/0]  
 | | | education = 5th-6th : <=50K (4.94/1.36) [2.15/0.36]  
 | | | education = 10th : <=50K (13.58/3.72) [7.07/3]  
 | | | education = 1st-4th : >50K (3.43/1.43) [0.36/0.36]  
 | | | education = Preschool : <=50K (0.72/0) [0/0]  
 | | | education = 12th : <=50K (4.43/1) [2/0]  
 | occupation = Handlers-cleaners  
 | | capital.gain < 5095.5  
 | | | education = Bachelors  
 | | | | fnlwgt < 235511 : <=50K (8.44/1.11) [8.35/4.11]  
 | | | | fnlwgt >= 235511 : >50K (3.11/0.03) [3.02/1.02]  
 | | | education = HS-grad  
 | | | | age < 44.5 : <=50K (87.35/10.02) [56.15/8]  
 | | | | age >= 44.5  
 | | | | | workclass = Self-emp-not-inc : <=50K (1.04/0.01) [0.02/0]  
 | | | | | workclass = Private  
 | | | | | | education.num < 9.26 : <=50K (28.14/13.25) [18.72/5.14]  
 | | | | | | education.num >= 9.26 : >50K (3/0) [2/0]  
 | | | | | | workclass = State-gov : <=50K (1.04/0.01) [1.02/1]  
 | | | | | | workclass = Federal-gov : <=50K (0/0) [0/0]

```

|| | | | workclass = Local-gov : <=50K (3.11/0.02) [1.07/0.01]
|| | | | workclass = Self-emp-inc : <=50K (0/0) [0/0]
|| | | | workclass = Without-pay : <=50K (0/0) [0/0]
|| | | | workclass = Never-worked : <=50K (0/0) [0/0]
|| | education = 11th : <=50K (12.2/1) [10.06/1]
|| | education = Masters : <=50K (2.24/0.11) [0.12/0.06]
|| | education = 9th : <=50K (12.08/1) [6.11/0]
|| | education = Some-college : <=50K (52.63/12.17) [27.35/5.11]
|| | education = Assoc-acdm : <=50K (5.08/0.02) [5.06/0.03]
|| | education = Assoc-voc : <=50K (13.14/1.11) [4.03/1]
|| | education = 7th-8th : <=50K (13.35/0.02) [7.08/0]
|| | education = Doctorate : >50K (0.09/0.02) [0.06/0.02]
|| | education = Prof-school : <=50K (0.09/0.03) [0.06/0.03]
|| | education = 5th-6th : <=50K (7.17/0.02) [8.08/0]
|| | education = 10th : <=50K (14.14/0.02) [9.05/1]
|| | education = 1st-4th : <=50K (7.06/0) [1.02/0]
|| | education = Preschool : <=50K (0.03/0) [0/0]
|| | education = 12th : <=50K (2.06/0) [2/2]
|| capital.gain >= 5095.5 : >50K (11.3/0.03) [4.11/0]
| occupation = Prof-specialty
|| education = Bachelors
|| | education.num < 13 : >50K (830/0) [397/0]
|| | education.num >= 13
|| | | education.num < 13
|| | | | capital.gain < 7032.5
|| | | | | workclass = Self-emp-not-inc : >50K (37.62/15.41) [17.39/4.28]
|| | | | | workclass = Private
|| | | | | age < 28.13 : <=50K (28.16/9) [15.47/5]
|| | | | | age >= 28.13
|| | | | | native.country = United-States
|| | | | | capital.gain < 4890.03
|| | | | | capital.loss < 863.42
|| | | | | house.per.week < 35.4 : <=50K (11.06/5.32) [12.21/5.63]
|| | | | | house.per.week >= 35.4 : >50K (278.11/35.51) [150.75/18.32]
|| | | | | capital.loss >= 863.42 : >50K (33.63/0) [14.16/1.16]
|| | | | | capital.gain >= 4890.03 : <=50K (3.32/0.16) [1/0]
|| | | | | native.country = Cuba : >50K (0/0) [0/0]
|| | | | | native.country = Jamaica : >50K (0/0) [0/0]
|| | | | | native.country = India : <=50K (1.02/0.02) [0/0]
|| | | | | native.country = Mexico : >50K (0/0) [0/0]
|| | | | | native.country = South : <=50K (0/0) [0.16/0]

```

							native.country =	Puerto-Rico : >50K (0/0) [0/0]
							native.country =	Honduras : >50K (0/0) [0/0]
							native.country =	England : >50K (1.02/0.01) [0/0]
							native.country =	Canada : <=50K (0/0) [2/1]
							native.country =	Germany : <=50K (1.02/0.02) [2/1]
							native.country =	Iran : >50K (1.02/1.01) [1/0]
							native.country =	Philippines : >50K (1.02/0.01) [1/0]
							native.country =	Italy : >50K (2.2/1.01) [1.01/0]
							native.country =	Poland : >50K (0/0) [0/0]
							native.country =	Columbia : <=50K (1.02/0.02) [0/0]
							native.country =	Cambodia : >50K (0/0) [0/0]
							native.country =	Thailand : >50K (0/0) [0/0]
							native.country =	Ecuador : <=50K (1.02/0.02) [0/0]
							native.country =	Laos : >50K (0/0) [0/0]
							native.country =	Taiwan : >50K (0/0) [0/0]
							native.country =	Haiti : >50K (0/0) [0/0]
							native.country =	Portugal : >50K (0/0) [0/0]
							native.country =	Dominican-Republic : >50K (0/0) [0/0]
							native.country =	El-Salvador : >50K (0/0) [0/0]
							native.country =	France : >50K (0/0) [0/0]
							native.country =	Guatemala : >50K (0/0) [0/0]
							native.country =	China : <=50K (0.16/0) [1/0]
							native.country =	Japan : >50K (1.02/0.01) [1/1]
							native.country =	Yugoslavia : >50K (0/0) [0/0]
							native.country =	Peru : >50K (0/0) [1/0]
							native.country =	Outlying-US(Guam-USVI-etc) : >50K (0/0) [0/0]
							native.country =	Scotland : >50K (0/0) [0/0]
							native.country =	Trinidad&Tobago : >50K (0/0) [0/0]
							native.country =	Greece : >50K (0/0) [0/0]
							native.country =	Nicaragua : >50K (0/0) [0/0]
							native.country =	Vietnam : <=50K (0.16/0) [0/0]
							native.country =	Hong : >50K (0/0) [0/0]
							native.country =	Ireland : >50K (2.04/0.01) [0.01/0]
							native.country =	Hungary : >50K (0/0) [0/0]
							native.country =	Holand-Netherlands : >50K (0/0) [0/0]
							workclass = State-gov	
							age < 51 : >50K (16.06/4.04) [15.05/6.04]	
							age >= 51 : <=50K (2.24/0.08) [1.14/0.05]	
							workclass = Federal-gov : >50K (17.28/6.19) [14.18/3.13]	
							workclass = Local-gov	
							age < 46.5	

```

| | | | | || | house.per.week < 32.5 : <=50K (2.07/0) [0.02/0]
| | | | | || | house.per.week >= 32.5
| | | | | || | fnlwgt < 207088.5
| | | | | || | fnlwgt < 161012.5
| | | | | || | | age < 34.5 : <=50K (8.02/3) [3.05/0]
| | | | | || | | age >= 34.5 : >50K (14/2) [4/1]
| | | | | || | | fnlwgt >= 161012.5 : <=50K (7.02/2) [4/3]
| | | | | || | | fnlwgt >= 207088.5 : >50K (9/1) [5/3]
| | | | | || | age >= 46.5 : <=50K (15.8/4.31) [9.51/4.17]
| | | | | | workclass = Self-emp-inc : >50K (12.2/5.13) [5.13/1.09]
| | | | | | workclass = Without-pay : >50K (0/0) [0/0]
| || | | | workclass = Never-worked : >50K (0/0) [0/0]
| || | | capital.gain >= 7032.5 : >50K (91.1/1) [40.44/0]
| || | education.num >= 13 : >50K (494/0) [229/0]
| | education = HS-grad
| || education.num < 9
| || | house.per.week < 32.79 : <=50K (21.29/4.44) [8.6/1.22]
| || | house.per.week >= 32.79
| || | education.num < 8.92 : >50K (16/0) [5/0]
| || | education.num >= 8.92
| || | | fnlwgt < 375725.43 : >50K (108.92/28.98) [80.42/21.23]
| || | | fnlwgt >= 375725.43 : <=50K (4.53/0.22) [0.44/0]
| || education.num >= 9 : >50K (526/0) [238/0]
| | education = 11th : <=50K (3.85/0) [1.88/0]
| | education = Masters
| || education.num < 13.99 : >50K (163/0) [98/0]
| || education.num >= 13.99
| || | capital.gain < 5095.5
| || | education.num < 14.01
| || | | house.per.week < 20.42 : <=50K (10.31/1.44) [5.44/1.22]
| || | | house.per.week >= 20.42
| || | | capital.loss < 931.95
| || | | | age < 32.5 : <=50K (27.22/12.22) [8/3]
| || | | | age >= 32.5 : >50K (186.97/44.1) [79.31/15.66]
| || | | | capital.loss >= 931.95 : >50K (33/1) [12/0]
| || | | education.num >= 14.01 : >50K (36/0) [15/0]
| || | capital.gain >= 5095.5 : >50K (72.22/0) [26.22/0]
| | education = 9th : <=50K (2.31/0.22) [1.53/0]
| | education = Some-college
| || education.num < 10.01
| || | education.num < 9.98 : >50K (51/0) [23/0]

```

| | | education.num >= 9.98  
 | | | capital.gain < 3909  
 | | | | house.per.week < 33.5 : <=50K (13.48/1.88) [8.75/1.44]  
 | | | | house.per.week >= 33.5  
 | | | | | race = White  
 | | | | | workclass = Self-emp-not-inc : >50K (7.22/2.15) [3.19/1.15]  
 | | | | | workclass = Private : >50K (86.65/28.77) [52.3/17.77]  
 | | | | | workclass = State-gov : <=50K (4.13/2.04) [1.11/0.03]  
 | | | | | workclass = Federal-gov : >50K (6.19/1.13) [1.16/0.13]  
 | | | | | workclass = Local-gov : <=50K (1.03/0.01) [3.03/1.01]  
 | | | | | workclass = Self-emp-inc : <=50K (2.06/0.02) [0.05/0.01]  
 | | | | | workclass = Without-pay : >50K (0/0) [0/0]  
 | | | | | workclass = Never-worked : >50K (0/0) [0/0]  
 | | | | | race = Black : >50K (4/1) [1.22/1]  
 | | | | | race = Asian-Pac-Islander : >50K (1.44/0) [0.22/0]  
 | | | | | race = Amer-Indian-Eskimo : <=50K (5/1) [0/0]  
 | | | | | race = Other : >50K (1/0) [0/0]  
 | | | capital.gain >= 3909 : >50K (19.66/-0) [7.44/1]  
 | | education.num >= 10.01 : >50K (197/0) [91/0]  
 | | education = Assoc-acdm : >50K (37.31/16.88) [11.88/4.44]  
 | | education = Assoc-voc  
 | | capital.gain < 3898.5  
 | | | workclass = Self-emp-not-inc : <=50K (5.25/0.19) [1.05/0]  
 | | | workclass = Private : >50K (32.53/12.34) [10.34/1.34]  
 | | | workclass = State-gov : <=50K (2.1/0.08) [1.02/1]  
 | | | workclass = Federal-gov : <=50K (1.05/0.04) [0.01/0]  
 | | | workclass = Local-gov : <=50K (1.05/0.04) [0.01/0]  
 | | | workclass = Self-emp-inc : >50K (0/0) [0/0]  
 | | | workclass = Without-pay : >50K (0/0) [0/0]  
 | | | workclass = Never-worked : >50K (0/0) [0/0]  
 | | capital.gain >= 3898.5 : >50K (11/0) [3/0]  
 | | education = 7th-8th : <=50K (10.04/0.22) [2.31/0.22]  
 | | education = Doctorate : >50K (174.75/18.22) [107.88/14.22]  
 | | education = Prof-school : >50K (353.75/26.88) [163.88/8.44]  
 | | education = 5th-6th : <=50K (3.41/0.22) [1.31/0.22]  
 | | education = 10th : <=50K (3.19/1.44) [0.66/0]  
 | | education = 1st-4th : <=50K (0.88/0) [2.22/0]  
 | | education = Preschool : <=50K (1.44/0) [0/0]  
 | | education = 12th : >50K (2.88/1.88) [1/0]  
 | occupation = Other-service  
 | education = Bachelors



|| | native.country = United-States  
 || | | age < 28.5 : <=50K (5.02/0) [1.06/1]  
 || | | age >= 28.5  
 || | | | age < 37.5 : >50K (8.06/1.06) [4/1]  
 || | | | age >= 37.5 : <=50K (7.73/1.34) [4.19/1.18]  
 || | native.country = Cuba : <=50K (1/0) [0.04/0]  
 || | native.country = Jamaica : <=50K (0/0) [0/0]  
 || | native.country = India : <=50K (0/0) [0/0]  
 || | native.country = Mexico : <=50K (1/0) [0.04/0]  
 || | native.country = South : <=50K (1/0) [0.06/0]  
 || | native.country = Puerto-Rico : <=50K (0/0) [0/0]  
 || | native.country = Honduras : <=50K (0/0) [0/0]  
 || | native.country = England : <=50K (0/0) [0/0]  
 || | native.country = Canada : <=50K (0/0) [0/0]  
 || | native.country = Germany : <=50K (0/0) [0/0]  
 || | native.country = Iran : >50K (0/0) [1/0]  
 || | native.country = Philippines : >50K (1/0) [0.04/0.04]  
 || | native.country = Italy : >50K (0.02/0) [0/0]  
 || | native.country = Poland : <=50K (0/0) [0/0]  
 || | native.country = Columbia : <=50K (0/0) [0/0]  
 || | native.country = Cambodia : <=50K (0/0) [0/0]  
 || | native.country = Thailand : <=50K (0/0) [0/0]  
 || | native.country = Ecuador : <=50K (0/0) [0/0]  
 || | native.country = Laos : <=50K (0/0) [0/0]  
 || | native.country = Taiwan : <=50K (0/0) [0/0]  
 || | native.country = Haiti : <=50K (0/0) [0/0]  
 || | native.country = Portugal : <=50K (0/0) [0/0]  
 || | native.country = Dominican-Republic : <=50K (0/0) [0/0]  
 || | native.country = El-Salvador : <=50K (0/0) [1/0]  
 || | native.country = France : <=50K (0/0) [0/0]  
 || | native.country = Guatemala : <=50K (0/0) [0/0]  
 || | native.country = China : <=50K (1.02/0) [0.04/0]  
 || | native.country = Japan : <=50K (0/0) [0/0]  
 || | native.country = Yugoslavia : <=50K (0/0) [1/0]  
 || | native.country = Peru : <=50K (0/0) [0/0]  
 || | native.country = Outlying-US(Guam-USVI-etc) : <=50K (0/0) [0/0]  
 || | native.country = Scotland : <=50K (0/0) [0/0]  
 || | native.country = Trinidad&Tobago : <=50K (0/0) [0/0]  
 || | native.country = Greece : <=50K (0/0) [0/0]  
 || | native.country = Nicaragua : <=50K (0/0) [0/0]  
 || | native.country = Vietnam : <=50K (2.02/0) [0.07/0]

| | | native.country = Hong : <=50K (0/0) [0/0]  
 | | | native.country = Ireland : <=50K (0/0) [0/0]  
 | | | native.country = Hungary : <=50K (0/0) [0/0]  
 | | | native.country = Holand-Netherlands : <=50K (0/0) [0/0]  
 | | education = HS-grad  
 | | | native.country = United-States  
 | | | capital.gain < 7032.5  
 | | | | age < 35.5 : <=50K (42.03/2) [15.06/1]  
 | | | | age >= 35.5  
 | | | | | house.per.week < 34.5 : <=50K (13.79/0.02) [6.39/1.02]  
 | | | | | house.per.week >= 34.5  
 | | | | | | house.per.week < 39.87 : >50K (4.08/0.06) [2.02/0.02]  
 | | | | | | house.per.week >= 39.87 : <=50K (74.68/17.34) [27.58/4.2]  
 | | | capital.gain >= 7032.5 : >50K (4/0) [2/0]  
 | | | native.country = Cuba : <=50K (1.02/0) [1.03/0]  
 | | | native.country = Jamaica : <=50K (0/0) [0/0]  
 | | | native.country = India : <=50K (0/0) [0/0]  
 | | | native.country = Mexico : <=50K (8.15/1) [3.07/1]  
 | | | native.country = South : <=50K (1.02/0) [0.01/0]  
 | | | native.country = Puerto-Rico : <=50K (0/0) [0.02/0]  
 | | | native.country = Honduras : <=50K (0/0) [0/0]  
 | | | native.country = England : <=50K (1.02/0) [0.01/0]  
 | | | native.country = Canada : >50K (0.06/0.02) [1/0]  
 | | | native.country = Germany : <=50K (1.02/1) [1.03/0]  
 | | | native.country = Iran : <=50K (0/0) [0/0]  
 | | | native.country = Philippines : <=50K (3.08/0) [0.02/0]  
 | | | native.country = Italy : <=50K (1.02/0) [0.01/0]  
 | | | native.country = Poland : <=50K (1.02/0) [0.01/0]  
 | | | native.country = Columbia : <=50K (0.02/0) [0/0]  
 | | | native.country = Cambodia : <=50K (0/0) [0/0]  
 | | | native.country = Thailand : <=50K (1.02/0) [0.01/0]  
 | | | native.country = Ecuador : <=50K (0/0) [0/0]  
 | | | native.country = Laos : <=50K (0/0) [0/0]  
 | | | native.country = Taiwan : <=50K (0/0) [0/0]  
 | | | native.country = Haiti : <=50K (2.04/0) [0.01/0]  
 | | | native.country = Portugal : <=50K (0/0) [0/0]  
 | | | native.country = Dominican-Republic : <=50K (0.04/0) [0/0]  
 | | | native.country = El-Salvador : >50K (1.02/0.02) [0.01/0.01]  
 | | | native.country = France : <=50K (0/0) [0/0]  
 | | | native.country = Guatemala : <=50K (1.02/0) [1.01/0]  
 | | | native.country = China : <=50K (5.09/0) [1.05/0]

| | native.country = Japan : <=50K (2.04/1) [0.01/0]  
 | | native.country = Yugoslavia : <=50K (1.02/1) [1.01/0]  
 | | native.country = Peru : <=50K (0/0) [0/0]  
 | | native.country = Outlying-US(Guam-USVI-etc) : <=50K (0/0) [0/0]  
 | | native.country = Scotland : <=50K (0/0) [0/0]  
 | | native.country = Trinidad&Tobago : <=50K (0/0) [0/0]  
 | | native.country = Greece : <=50K (0/0) [0/0]  
 | | native.country = Nicaragua : <=50K (1.02/0) [0.01/0]  
 | | native.country = Vietnam : <=50K (0/0) [0/0]  
 | | native.country = Hong : <=50K (0/0) [0/0]  
 | | native.country = Ireland : <=50K (0/0) [1/0]  
 | | native.country = Hungary : <=50K (0/0) [0/0]  
 | | native.country = Holand-Netherlands : <=50K (0/0) [0/0]  
 | | education = 11th : <=50K (13.26/1) [11.08/1]  
 | | education = Masters : <=50K (4.34/0.16) [2.18/1.1]  
 | | education = 9th : <=50K (19.12/0.02) [4.14/1]  
 | | education = Some-college  
 | | education.num < 9.5 : >50K (4/0) [3/0]  
 | | education.num >= 9.5  
 | | | education.num < 10.24 : <=50K (59.91/8.28) [33.51/8.18]  
 | | | education.num >= 10.24 : >50K (3/0) [3/0]  
 | | education = Assoc-acdm : <=50K (6.12/0.04) [3.08/2.04]  
 | | education = Assoc-voc : <=50K (19.18/4.14) [6.04/2]  
 | | education = 7th-8th : <=50K (13.47/0.02) [3.12/0.02]  
 | | education = Doctorate : >50K (0.16/0.02) [0.08/0.02]  
 | | education = Prof-school : <=50K (1.16/0.08) [0.08/0.04]  
 | | education = 5th-6th : <=50K (15.22/0.02) [10.12/0.02]  
 | | education = 10th : <=50K (14.2/0.04) [7.06/1]  
 | | education = 1st-4th : <=50K (8.08/0) [3.02/0]  
 | | education = Preschool : <=50K (3.04/0) [1/0]  
 | | education = 12th : <=50K (4.08/1) [3/0]  
 | occupation = Adm-clerical  
 | | education = Bachelors  
 | | | education.num < 12.99 : >50K (79/0) [35/0]  
 | | | education.num >= 12.99  
 | | | | native.country = United-States  
 | | | | capital.gain < 5095.5  
 | | | | | house.per.week < 39 : <=50K (4.49/0.13) [6.49/3.13]  
 | | | | | house.per.week >= 39  
 | | | | | workclass = Self-emp-not-inc : >50K (1.01/0.01) [0/0]  
 | | | | | workclass = Private

```

| | | | | | | | | | age < 38.21
| | | | | | | | | | capital.loss < 1733 : <=50K (11.02/5) [5.88/2.86]
| | | | | | | | | | capital.loss >= 1733 : >50K (3/0) [1.02/0.02]
| | | | | | | | | | age >= 38.21
| | | | | | | | | | fnlwgt < 214364
| | | | | | | | | | fnlwgt < 167403.24
| | | | | | | | | | age < 55.5 : >50K (19.88/1.02) [7.86/0]
| | | | | | | | | | age >= 55.5 : <=50K (2.09/1) [1.07/0.02]
| | | | | | | | | | fnlwgt >= 167403.24 : <=50K (3.14/1.07) [1.07/0.05]
| | | | | | | | | | fnlwgt >= 214364 : >50K (11/0.05) [4/3]
| | | | | | | | | | workclass = State-gov
| | | | | | | | | | age < 41.5 : <=50K (2/0) [2.01/1]
| | | | | | | | | | age >= 41.5 : >50K (5.06/2.03) [1.02/0.01]
| | | | | | | | | | workclass = Federal-gov : >50K (4.03/0.02) [4.02/2.01]
| | | | | | | | | | workclass = Local-gov
| | | | | | | | | | age < 29.5 : <=50K (2/0) [0.01/0]
| | | | | | | | | | age >= 29.5 : >50K (5.06/0.03) [1.02/1.01]
| | | | | | | | | | workclass = Self-emp-inc : <=50K (1.01/1) [1/0]
| | | | | | | | | | workclass = Without-pay : >50K (0/0) [0/0]
| | | | | | | | | | workclass = Never-worked : >50K (0/0) [0/0]
| | | | | | | | | | capital.gain >= 5095.5 : >50K (19.23/0.03) [8.07/0]
| | | | | | | | | | native.country = Cuba : >50K (0/0) [0/0]
| | | | | | | | | | native.country = Jamaica : >50K (0/0) [0/0]
| | | | | | | | | | native.country = India : <=50K (1.02/0.02) [0.02/0.02]
| | | | | | | | | | native.country = Mexico : <=50K (1.02/0.02) [1.02/0.02]
| | | | | | | | | | native.country = South : <=50K (0/0) [0.03/0]
| | | | | | | | | | native.country = Puerto-Rico : >50K (0/0) [0/0]
| | | | | | | | | | native.country = Honduras : >50K (0/0) [0/0]
| | | | | | | | | | native.country = England : <=50K (0/0) [1/0]
| | | | | | | | | | native.country = Canada : <=50K (1.02/0.02) [0.02/0.02]
| | | | | | | | | | native.country = Germany : >50K (1.02/0) [0.02/0]
| | | | | | | | | | native.country = Iran : >50K (1.02/0) [0.02/0]
| | | | | | | | | | native.country = Philippines : >50K (6.11/2) [4.11/1]
| | | | | | | | | | native.country = Italy : <=50K (1.05/0.05) [0.02/0.02]
| | | | | | | | | | native.country = Poland : >50K (0/0) [0/0]
| | | | | | | | | | native.country = Columbia : >50K (0/0) [0/0]
| | | | | | | | | | native.country = Cambodia : >50K (0/0) [0/0]
| | | | | | | | | | native.country = Thailand : >50K (0/0) [0/0]
| | | | | | | | | | native.country = Ecuador : >50K (0/0) [0/0]
| | | | | | | | | | native.country = Laos : >50K (0/0) [0/0]
| | | | | | | | | | native.country = Taiwan : >50K (0/0) [0/0]

```

```

| | | native.country = Haiti : >50K (0/0) [0/0]
| | | native.country = Portugal : >50K (0/0) [0/0]
| | | native.country = Dominican-Republic : >50K (0/0) [0/0]
| | | native.country = El-Salvador : >50K (0/0) [0/0]
| | | native.country = France : >50K (0/0) [0/0]
| | | native.country = Guatemala : >50K (0/0) [0/0]
| | | native.country = China : <=50K (1.05/0.02) [0.02/0.02]
| | | native.country = Japan : <=50K (1.02/0.02) [0.02/0.02]
| | | native.country = Yugoslavia : >50K (0/0) [0/0]
| | | native.country = Peru : >50K (0/0) [0/0]
| | | native.country = Outlying-US(Guam-USVI-etc) : >50K (0/0) [0/0]
| | | native.country = Scotland : >50K (0/0) [0/0]
| | | native.country = Trinidad&Tobago : >50K (0/0) [0/0]
| | | native.country = Greece : >50K (0/0) [0/0]
| | | native.country = Nicaragua : >50K (0/0) [0/0]
| | | native.country = Vietnam : <=50K (1.05/0.02) [0.02/0.02]
| | | native.country = Hong : >50K (0/0) [0/0]
| | | native.country = Ireland : >50K (0/0) [0/0]
| | | native.country = Hungary : >50K (0/0) [0/0]
| | | native.country = Holand-Netherlands : >50K (0/0) [0/0]
| | education = HS-grad
| | education.num < 9.01
| | age < 37.5
| | | fnlwgt < 41113.72 : >50K (2/0) [2/1]
| | | fnlwgt >= 41113.72
| | | | house.per.week < 47.5 : <=50K (34.46/2.03) [21.23/5]
| | | | house.per.week >= 47.5
| | | | | age < 32 : <=50K (4/0) [1/0]
| | | | | age >= 32 : >50K (4.07/0.07) [1/0]
| | | age >= 37.5
| | | | house.per.week < 26.9 : <=50K (8.15/0.03) [4.46/0]
| | | | house.per.week >= 26.9
| | | | | capital.gain < 3500.29
| | | | | age < 49.86
| | | | | | age < 41.03 : <=50K (14.03/5) [10.03/3]
| | | | | | age >= 41.03
| | | | | | | workclass = Self-emp-not-inc : >50K (0/0) [0/0]
| | | | | | | workclass = Private
| | | | | | | | house.per.week < 39 : <=50K (2.02/0) [1.02/0]
| | | | | | | | house.per.week >= 39
| | | | | | | | fnlwgt < 104328.5 : >50K (4/0) [2/2]

```

```
| | | | | | | | | | fnlwgt >= 104328.5  
| | | | | | | | | | fnlwgt < 115937 : <=50K (2/0) [1/0]  
| | | | | | | | | | fnlwgt >= 115937 : >50K (16.21/4.19) [2.09/0.09]  
  
| | | | | | | | workclass = State-gov : <=50K (1.01/1) [1/0]  
| | | | | | | | workclass = Federal-gov : >50K (9.09/0.08) [6.04/3.04]  
| | | | | | | | workclass = Local-gov : >50K (0/0) [0/0]  
| | | | | | | | workclass = Self-emp-inc : >50K (0/0) [0/0]  
| | | | | | | | workclass = Without-pay : >50K (0/0) [0/0]  
| | | | | | | | workclass = Never-worked : >50K (0/0) [0/0]  
  
| | | | | | age >= 49.86  
  
| | | | | education.num < 8.71 : >50K (2/0) [0/0]  
| | | | | education.num >= 8.71  
  
| | | | | | | | workclass = Self-emp-not-inc : <=50K (0/0) [0/0]  
| | | | | | | | workclass = Private : <=50K (14.7/3.24) [14.55/6.16]  
| | | | | | | | workclass = State-gov : >50K (3.15/1.1) [0.12/0.08]  
| | | | | | | | workclass = Federal-gov  
| | | | | | | | age < 61.5 : >50K (9.33/4.21) [5.23/1.15]  
| | | | | | | | age >= 61.5 : <=50K (4.32/0.11) [1.28/1.07]  
| | | | | | | | workclass = Local-gov : <=50K (1.05/0.02) [1.04/0.01]  
| | | | | | | | workclass = Self-emp-inc : <=50K (1.05/0.02) [0.04/0.01]  
| | | | | | | | workclass = Without-pay : <=50K (0/0) [0/0]  
| | | | | | | | workclass = Never-worked : <=50K (0/0) [0/0]  
  
| | | capital.gain >= 3500.29 : >50K (8.07/0.03) [6.03/1.03]  
| || education.num >= 9.01 : >50K (61/-0) [29/0]  
  
| | education = 11th : <=50K (5.43/2) [3.13/0]  
| | education = Masters : >50K (21.56/8.3) [6.3/3.13]  
| | education = 9th : <=50K (2.2/0.03) [1.23/1]  
| | education = Some-college  
  
|| capital.gain < 5095.5  
|| education.num < 10.01  
|| native.country = United-States  
|| age < 37.5 : <=50K (46.13/16) [17.07/5]  
|| age >= 37.5  
  
|| house.per.week < 40.59  
  
|| workclass = Self-emp-not-inc : <=50K (0/0) [0/0]  
|| workclass = Private : <=50K (22.53/9.14) [15.28/5.09]  
|| workclass = State-gov : <=50K (4.04/1.03) [1.05/0.02]  
|| workclass = Federal-gov  
|| fnlwgt < 135271.5 : <=50K (2.1/0.01) [1.06/0.01]  
|| fnlwgt >= 135271.5 : >50K (10.19/1.13) [4.1/2.06]  
|| workclass = Local-gov : >50K (3.07/1.05) [0.04/0.03]
```

| | | | | | | workclass = Self-emp-inc : <=50K (0/0) [0/0]  
 | | | | | | | workclass = Without-pay : <=50K (0/0) [0/0]  
 | | | | | | | workclass = Never-worked : <=50K (0/0) [0/0]  
 | | | | | | | house.per.week >= 40.59  
 | | | | | | | workclass = Self-emp-not-inc : >50K (1.01/0) [0.01/0.01]  
 | | | | | | | workclass = Private : >50K (11.07/0.05) [6.12/0.07]  
 | | | | | | | workclass = State-gov : >50K (0/0) [0/0]  
 | | | | | | | workclass = Federal-gov : <=50K (2.01/1) [0.02/0.01]  
 | | | | | | | workclass = Local-gov : <=50K (1.01/0) [0.01/0]  
 | | | | | | | workclass = Self-emp-inc : >50K (0/0) [0/0]  
 | | | | | | | workclass = Without-pay : >50K (0/0) [0/0]  
 | | | | | | | workclass = Never-worked : >50K (0/0) [0/0]  
 | | | | | native.country = Cuba : <=50K (0/0) [0/0]  
 | | | | | native.country = Jamaica : >50K (0/0) [1/0]  
 | | | | | native.country = India : <=50K (0/0) [1/0]  
 | | | | | native.country = Mexico : <=50K (1.04/0) [2/0]  
 | | | | | native.country = South : >50K (0.03/0) [0/0]  
 | | | | | native.country = Puerto-Rico : >50K (1.01/0.01) [0/0]  
 | | | | | native.country = Honduras : <=50K (0/0) [0/0]  
 | | | | | native.country = England : <=50K (0/0) [0/0]  
 | | | | | native.country = Canada : >50K (0.03/0.03) [1/0]  
 | | | | | native.country = Germany : <=50K (1.01/0) [0/0]  
 | | | | | native.country = Iran : <=50K (0/0) [0/0]  
 | | | | | native.country = Philippines : >50K (2.05/0.02) [0/0]  
 | | | | | native.country = Italy : <=50K (0/0) [0/0]  
 | | | | | native.country = Poland : <=50K (0/0) [0/0]  
 | | | | | native.country = Columbia : <=50K (0/0) [0/0]  
 | | | | | native.country = Cambodia : <=50K (0/0) [0/0]  
 | | | | | native.country = Thailand : <=50K (0/0) [0/0]  
 | | | | | native.country = Ecuador : <=50K (0/0) [0/0]  
 | | | | | native.country = Laos : <=50K (0/0) [1/0]  
 | | | | | native.country = Taiwan : <=50K (0/0) [0/0]  
 | | | | | native.country = Haiti : <=50K (0/0) [0/0]  
 | | | | | native.country = Portugal : <=50K (0/0) [0/0]  
 | | | | | native.country = Dominican-Republic : <=50K (0/0) [0/0]  
 | | | | | native.country = El-Salvador : <=50K (0/0) [0/0]  
 | | | | | native.country = France : <=50K (0/0) [0/0]  
 | | | | | native.country = Guatemala : <=50K (0/0) [0/0]  
 | | | | | native.country = China : <=50K (0/0) [0/0]  
 | | | | | native.country = Japan : <=50K (0/0) [0/0]  
 | | | | | native.country = Yugoslavia : <=50K (0/0) [0/0]

|| | | | native.country = Peru : <=50K (0/0) [0/0]  
 || | | | native.country = Outlying-US(Guam-USVI-etc) : <=50K (0/0) [0/0]  
 || | | | native.country = Scotland : <=50K (0/0) [0/0]  
 || | | | native.country = Trinidad&Tobago : <=50K (0/0) [0/0]  
 || | | | native.country = Greece : <=50K (0/0) [0/0]  
 || | | | native.country = Nicaragua : <=50K (0.03/0) [0/0]  
 || | | | native.country = Vietnam : <=50K (1.01/0) [0/0]  
 || | | | native.country = Hong : <=50K (0/0) [0/0]  
 || | | | native.country = Ireland : <=50K (0/0) [0/0]  
 || | | | native.country = Hungary : <=50K (0/0) [0/0]  
 || | | | native.country = Holand-Netherlands : <=50K (0/0) [0/0]  
 || | | education.num >= 10.01 : >50K (8/-0) [7/0]  
 || | capital.gain >= 5095.5 : >50K (19.1/0) [10.07/0]  
 || education = Assoc-acdm : <=50K (18.2/9.07) [15.13/5.07]  
 || education = Assoc-voc : <=50K (27.3/10.23) [6.07/2]  
 || education = 7th-8th : <=50K (3.75/0.03) [2.2/0.03]  
 || education = Doctorate : >50K (1.26/0.03) [1.13/0.03]  
 || education = Prof-school : >50K (5.26/2.13) [1.13/0.07]  
 || education = 5th-6th : <=50K (2.36/0.03) [0.2/0.03]  
 || education = 10th : <=50K (6.33/0.07) [3.1/0]  
 || education = 1st-4th : <=50K (0.13/0) [0.03/0]  
 || education = Preschool : <=50K (0.07/0) [0/0]  
 || education = 12th : <=50K (1.13/0) [5/0]  
 | occupation = Sales  
 || education = Bachelors  
 || | education.num < 12.99 : >50K (287/0) [150/0]  
 || | education.num >= 12.99  
 || | | education.num < 13  
 || | | | capital.gain < 4460.45  
 || | | | | capital.loss < 1772.81  
 || | | | | | native.country = United-States  
 || | | | | | | house.per.week < 35.37 : <=50K (18.26/5.2) [13.45/6.4]  
 || | | | | | | house.per.week >= 35.37  
 || | | | | | | | workclass = Self-emp-not-inc : >50K (29.15/12.09) [11.03/4]  
 || | | | | | | | workclass = Private  
 || | | | | | | | | age < 33.26 : >50K (53.08/23.08) [22.08/7.08]  
 || | | | | | | | | age >= 33.26  
 || | | | | | | | | | fnlwgt < 123977  
 || | | | | | | | | | | age < 38.97 : >50K (18/1) [4/1]  
 || | | | | | | | | | | age >= 38.97  
 || | | | | | | | | | | | age < 58.46



| | | | | | | | | | | | | | house.per.week < 52.5 : >50K (18/8) [6/1]  
 | | | | | | | | | | | | | | house.per.week >= 52.5 : <=50K (8.08/1) [1/0]  
 | | | | | | | | | | | | | | age >= 58.46 : >50K (5.15/0.15) [1.15/0.08]  
 | | | | | | | | | | | | | | fnlwgt >= 123977 : >50K (97.93/11.54) [54.23/9.08]  
 | | | | | | | | | | | | | | workclass = State-gov : <=50K (1.01/0) [0/0]  
 | | | | | | | | | | | | | | workclass = Federal-gov : <=50K (1.01/0) [1/0]  
 | | | | | | | | | | | | | | workclass = Local-gov : >50K (1.01/0) [0/0]  
 | | | | | | | | | | | | | | workclass = Self-emp-inc  
 | | | | | | | | | | | | | | house.per.week < 47.5  
 | | | | | | | | | | | | | | age < 39.5 : <=50K (3.01/0) [1/1]  
 | | | | | | | | | | | | | | age >= 39.5 : >50K (5.08/1.09) [1.02/0.01]  
 | | | | | | | | | | | | | | house.per.week >= 47.5  
 | | | | | | | | | | | | | | fnlwgt < 100850 : <=50K (2.01/1) [1.01/0]  
 | | | | | | | | | | | | | | fnlwgt >= 100850 : >50K (14.99/0) [10.03/5.01]  
 | | | | | | | | | | | | | | workclass = Without-pay : >50K (0/0) [0/0]  
 | | | | | | | | | | | | | | workclass = Never-worked : >50K (0/0) [0/0]  
 | | | | | | | | | | | | | | native.country = Cuba : >50K (0/0) [0/0]  
 | | | | | | | | | | | | | | native.country = Jamaica : <=50K (0/0) [2/0]  
 | | | | | | | | | | | | | | native.country = India : <=50K (1.01/0.01) [2.01/1]  
 | | | | | | | | | | | | | | native.country = Mexico : >50K (1.01/0.01) [0.01/0.01]  
 | | | | | | | | | | | | | | native.country = South : <=50K (0/0) [2/1]  
 | | | | | | | | | | | | | | native.country = Puerto-Rico : <=50K (0/0) [1/0]  
 | | | | | | | | | | | | | | native.country = Honduras : >50K (0/0) [0/0]  
 | | | | | | | | | | | | | | native.country = England : >50K (0/0) [0/0]  
 | | | | | | | | | | | | | | native.country = Canada : >50K (3.04/0.02) [0.02/0.02]  
 | | | | | | | | | | | | | | native.country = Germany : >50K (0/0) [1/0]  
 | | | | | | | | | | | | | | native.country = Iran : >50K (0/0) [0/0]  
 | | | | | | | | | | | | | | native.country = Philippines : >50K (0/0) [0/0]  
 | | | | | | | | | | | | | | native.country = Italy : >50K (0.1/0) [0/0]  
 | | | | | | | | | | | | | | native.country = Poland : <=50K (1.01/0.01) [0.01/0]  
 | | | | | | | | | | | | | | native.country = Columbia : >50K (0/0) [0/0]  
 | | | | | | | | | | | | | | native.country = Cambodia : >50K (0/0) [0/0]  
 | | | | | | | | | | | | | | native.country = Thailand : >50K (0/0) [0/0]  
 | | | | | | | | | | | | | | native.country = Ecuador : >50K (0/0) [0/0]  
 | | | | | | | | | | | | | | native.country = Laos : >50K (0/0) [0/0]  
 | | | | | | | | | | | | | | native.country = Taiwan : >50K (0/0) [0/0]  
 | | | | | | | | | | | | | | native.country = Haiti : >50K (0/0) [0/0]  
 | | | | | | | | | | | | | | native.country = Portugal : >50K (0/0) [0/0]  
 | | | | | | | | | | | | | | native.country = Dominican-Republic : >50K (0/0) [0/0]  
 | | | | | | | | | | | | | | native.country = El-Salvador : >50K (0/0) [0/0]  
 | | | | | | | | | | | | | | native.country = France : >50K (0/0) [0/0]

```

| | | | | | native.country = Guatemala : >50K (0/0) [0/0]
| | | | | | native.country = China : <=50K (0.1/0) [0/0]
| | | | | | native.country = Japan : <=50K (1.01/0.01) [0.01/0]
| | | | | | native.country = Yugoslavia : >50K (0/0) [0/0]
| | | | | | native.country = Peru : <=50K (1.01/0.01) [0.01/0]
| | | | | | native.country = Outlying-US(Guam-USVI-etc) : >50K (0/0) [0/0]
| | | | | | native.country = Scotland : >50K (0/0) [0/0]
| | | | | | native.country = Trinidad&Tobago : >50K (0/0) [0/0]
| | | | | | native.country = Greece : >50K (0/0) [0/0]
| | | | | | native.country = Nicaragua : >50K (0/0) [0/0]
| | | | | | native.country = Vietnam : <=50K (1.11/0.01) [0.01/0]
| | | | | | native.country = Hong : >50K (0/0) [0/0]
| | | | | | native.country = Ireland : >50K (0/0) [0/0]
| | | | | | native.country = Hungary : >50K (0/0) [0/0]
| | | | | | native.country = Holand-Netherlands : >50K (0/0) [0/0]
| | | | | capital.loss >= 1772.81 : >50K (32.4/1) [17.3/0.3]
| | | | | capital.gain >= 4460.45 : >50K (45.69/0.1) [34.2/0]
| | | | | education.num >= 13 : >50K (91/0) [35/0]
| | | | | education = HS-grad
| | | | | education.num < 9
| | | | | capital.gain < 5095.5
| | | | | education.num < 8.99 : >50K (16/0) [6/0]
| | | | | education.num >= 8.99
| | | | | age < 31.05 : <=50K (46.5/5) [24.1/3]
| | | | | age >= 31.05
| | | | | age < 58.5
| | | | | capital.loss < 1760
| | | | | | | | workclass = Self-emp-not-inc : <=50K (32.59/9.09) [12.36/5.05]
| | | | | | | | workclass = Private : >50K (162.95/79.49) [66.81/30.59]
| | | | | | | | workclass = State-gov : <=50K (0/0) [0/0]
| | | | | | | | workclass = Federal-gov : <=50K (0/0) [0/0]
| | | | | | | | workclass = Local-gov : <=50K (0/0) [0/0]
| | | | | | | | workclass = Self-emp-inc : >50K (18.33/7.28) [9.2/4.18]
| | | | | | | | workclass = Without-pay : <=50K (0/0) [0/0]
| | | | | | | | workclass = Never-worked : <=50K (0/0) [0/0]
| | | | | | | capital.loss >= 1760 : >50K (16/2) [5/0]
| | | | | | | age >= 58.5 : <=50K (44.65/7.39) [33.07/8.79]
| | | | | | | capital.gain >= 5095.5 : >50K (24.2/0.1) [12/0]
| | | | | | | education.num >= 9 : >50K (207/-0) [98/0]
| | | | | | | education = 11th : <=50K (10.29/0) [8.4/3]
| | | | | | | education = Masters

```

```

| | capital.gain < 2665.5
| | | education.num < 14 : >50K (14/0) [4/0]
| | | education.num >= 14
| | | | workclass = Self-emp-not-inc
| | | | | house.per.week < 48 : >50K (2.21/0.12) [2.1/0.05]
| | | | | house.per.week >= 48 : <=50K (5.07/0.03) [1.03/0.02]
| | | | workclass = Private : >50K (19.75/8.42) [16.38/3.19]
| | | | workclass = State-gov : >50K (0/0) [0/0]
| | | | workclass = Federal-gov : >50K (0/0) [1/0]
| | | | workclass = Local-gov : >50K (0/0) [0/0]
| | | | workclass = Self-emp-inc : >50K (14.56/2.31) [5.28/1.14]
| | | | workclass = Without-pay : >50K (0/0) [0/0]
| | | | workclass = Never-worked : >50K (0/0) [0/0]
| | capital.gain >= 2665.5 : >50K (17.1/0) [10.1/2]
| | education = 9th : <=50K (4.6/2.1) [3.69/0]
| | education = Some-college
| | education.num < 10.03
| | | education.num < 9.99 : >50K (34/0) [17/0]
| | | education.num >= 9.99
| | | | capital.gain < 5095.5
| | | | | age < 33.5
| | | | | fnlwgt < 128461 : <=50K (14.2/0) [11.1/2]
| | | | | fnlwgt >= 128461
| | | | | | workclass = Self-emp-not-inc : >50K (2.01/0.01) [3.01/2.01]
| | | | | | workclass = Private
| | | | | | fnlwgt < 199871.5
| | | | | | fnlwgt < 181165.5
| | | | | | fnlwgt < 157294.5 : >50K (5/2) [1/0]
| | | | | | fnlwgt >= 157294.5 : <=50K (4/0) [2/0]
| | | | | | fnlwgt >= 181165.5 : >50K (4.09/1.09) [2/0]
| | | | | | fnlwgt >= 199871.5
| | | | | | house.per.week < 42
| | | | | | house.per.week < 36 : <=50K (2.09/0) [0/0]
| | | | | | house.per.week >= 36
| | | | | | age < 29 : <=50K (4/1) [4.09/1]
| | | | | | age >= 29 : >50K (2/0) [1/0]
| | | | | | house.per.week >= 42 : <=50K (12/0) [5/0]
| | | | | | workclass = State-gov : <=50K (0/0) [0/0]
| | | | | | workclass = Federal-gov : <=50K (0/0) [0/0]
| | | | | | workclass = Local-gov : <=50K (0/0) [0/0]
| | | | | | workclass = Self-emp-inc : <=50K (0/0) [0/0]

```

```

|| | | | | | | | workclass = Without-pay : <=50K (0/0) [0/0]
|| | | | | | | | workclass = Never-worked : <=50K (0/0) [0/0]
|| | | | | | | | age >= 33.5
|| | | | | | | | workclass = Self-emp-not-inc
|| | | | | | | | | fnlwgt < 300010 : <=50K (19.41/2.11) [10.23/2.07]
|| | | | | | | | | fnlwgt >= 300010 : >50K (3.05/0.02) [2.02/0.01]
|| | | | | | | | | workclass = Private
|| | | | | | | | | capital.gain < 3120
|| | | | | | | | | capital.loss < 1813.5
|| | | | | | | | | | fnlwgt < 121563
|| | | | | | | | | | | fnlwgt < 48819.5 : >50K (9.07/2.07) [3/2]
|| | | | | | | | | | | fnlwgt >= 48819.5
|| | | | | | | | | | | fnlwgt < 59806.5 : <=50K (6.07/0) [4.14/1.07]
|| | | | | | | | | | | fnlwgt >= 59806.5
|| | | | | | | | | | | | house.per.week < 32.5 : <=50K (3.21/0) [0.07/0]
|| | | | | | | | | | | | house.per.week >= 32.5 : >50K (22.07/11) [8.36/3.29]
|| | | | | | | | | | | | fnlwgt >= 121563
|| | | | | | | | | | | | fnlwgt < 199769.5 : >50K (40.93/8.64) [27.43/11.29]
|| | | | | | | | | | | | fnlwgt >= 199769.5 : >50K (40.14/17.71) [26.36/11.21]
|| | | | | | | | | | | | capital.loss >= 1813.5 : >50K (8.07/0.07) [4.07/2]
|| | | | | | | | | | | | capital.gain >= 3120 : <=50K (4.14/0) [3.07/0]
|| | | | | | | | | | | | workclass = State-gov : >50K (1.02/0.01) [0.01/0.01]
|| | | | | | | | | | | | workclass = Federal-gov : >50K (0/0) [0/0]
|| | | | | | | | | | | | workclass = Local-gov : <=50K (1.02/0.01) [0.01/0]
|| | | | | | | | | | | | workclass = Self-emp-inc : >50K (27.56/9.4) [15.31/6.21]
|| | | | | | | | | | | | workclass = Without-pay : >50K (0/0) [0/0]
|| | | | | | | | | | | | workclass = Never-worked : >50K (0/0) [0/0]
|| | | | | | | | | | | | capital.gain >= 5095.5 : >50K (19.3/0) [15.2/1]
|| | | | | | | | | | | | education.num >= 10.03 : >50K (59/-0) [32/0]
|| | | | | | | | | | | | education = Assoc-acdm
|| | | | | | | | | | | | capital.gain < 1551.5
|| | | | | | | | | | | | capital.loss < 943.5
|| | | | | | | | | | | | fnlwgt < 177267.5
|| | | | | | | | | | | | fnlwgt < 127268 : <=50K (11.2/3.1) [6/1]
|| | | | | | | | | | | | fnlwgt >= 127268 : >50K (4/0) [2.1/1]
|| | | | | | | | | | | | fnlwgt >= 177267.5 : <=50K (10.3/1) [11.3/5.1]
|| | | | | | | | | | | | capital.loss >= 943.5 : >50K (3/0) [3/0]
|| | | | | | | | | | | | capital.gain >= 1551.5 : >50K (14.1/0) [7/0]
|| | | | | | | | | | | | education = Assoc-voc : >50K (46.89/21.2) [18.2/8.2]
|| | | | | | | | | | | | education = 7th-8th : <=50K (15.28/4.1) [4.6/0.1]
|| | | | | | | | | | | | education = Doctorate : >50K (2.79/1.1) [3.4/1.1]

```

```

| | education = Prof-school : >50K (12.79/2.4) [9.4/2.2]
| | education = 5th-6th : <=50K (5.09/1.1) [2.6/2.1]
| | education = 10th : <=50K (11.99/1.2) [7.3/2]
| | education = 1st-4th : <=50K (4.4/0) [0.1/0]
| | education = Preschool : <=50K (0.2/0) [0/0]
| | education = 12th : >50K (4.4/0.4) [1/1]
| occupation = Craft-repair
| | education = Bachelors
| | | education.num < 13 : >50K (189/0) [79/0]
| | | education.num >= 13
| | | | education.num < 13.01
| | | | | house.per.week < 40.49
| | | | | workclass = Self-emp-not-inc : <=50K (8.66/1.22) [3.36/2.11]
| | | | | workclass = Private : >50K (44.36/18.24) [25.87/10.31]
| | | | | workclass = State-gov : >50K (0/0) [0/0]
| | | | | workclass = Federal-gov : <=50K (1.08/1.03) [1.05/0.01]
| | | | | workclass = Local-gov : <=50K (6.49/1.16) [1.27/0.08]
| | | | | workclass = Self-emp-inc : <=50K (1.08/0.03) [1.05/0.01]
| | | | | workclass = Without-pay : >50K (0/0) [0/0]
| | | | | workclass = Never-worked : >50K (0/0) [0/0]
| | | | house.per.week >= 40.49
| | | | age < 33.5
| | | | | workclass = Self-emp-not-inc : <=50K (2/0) [0.02/0]
| | | | | workclass = Private
| | | | | fnlwgt < 137780 : <=50K (2/0) [1.09/0]
| | | | | fnlwgt >= 137780 : >50K (7/1) [4/1]
| | | | | workclass = State-gov : >50K (0/0) [0/0]
| | | | | workclass = Federal-gov : >50K (0/0) [0/0]
| | | | | workclass = Local-gov : >50K (0/0) [0/0]
| | | | | workclass = Self-emp-inc : <=50K (2/1) [0.02/0]
| | | | | workclass = Without-pay : >50K (0/0) [0/0]
| | | | | workclass = Never-worked : >50K (0/0) [0/0]
| | | | age >= 33.5 : >50K (50.91/5.13) [25.65/10.26]
| | | education.num >= 13.01 : >50K (29/-0) [12/0]
| | education = HS-grad
| | | education.num < 9
| | | capital.gain < 5095.5
| | | | education.num < 8.97 : >50K (31/0) [12/0]
| | | | education.num >= 8.97
| | | | | house.per.week < 39 : <=50K (61.36/1.26) [36.86/4.13]
| | | | | house.per.week >= 39

```

[illegible]

[illegible]

| | | | | | | native.country = Laos : <=50K (0/0) [0/0]  
 | | | | | | | native.country = Taiwan : <=50K (1.01/0) [0.01/0]  
 | | | | | | | native.country = Haiti : <=50K (1.01/0) [0.01/0]  
 | | | | | | | native.country = Portugal : <=50K (1.01/0) [0.01/0]  
 | | | | | | | native.country = Dominican-Republic : <=50K (0.13/0) [0/0]  
 | | | | | | | native.country = El-Salvador : <=50K (1.01/1) [1.01/0]  
 | | | | | | | native.country = France : <=50K (0/0) [0/0]  
 | | | | | | | native.country = Guatemala : <=50K (0/0) [0/0]  
 | | | | | | | native.country = China : <=50K (0/0) [0/0]  
 | | | | | | | native.country = Japan : <=50K (1.01/0) [0.01/0]  
 | | | | | | | native.country = Yugoslavia : <=50K (0/0) [0/0]  
 | | | | | | | native.country = Peru : <=50K (0/0) [0/0]  
 | | | | | | | native.country = Outlying-US(Guam-USVI-etc) : <=50K (0/0) [0/0]  
 | | | | | | | native.country = Scotland : <=50K (0/0) [0/0]  
 | | | | | | | native.country = Trinidad&Tobago : <=50K (0/0) [0/0]  
 | | | | | | | native.country = Greece : <=50K (1.01/0) [1.01/0]  
 | | | | | | | native.country = Nicaragua : <=50K (0/0) [0/0]  
 | | | | | | | native.country = Vietnam : <=50K (1.01/0) [0.01/0]  
 | | | | | | | native.country = Hong : <=50K (0/0) [0/0]  
 | | | | | | | native.country = Ireland : <=50K (0/0) [0/0]  
 | | | | | | | native.country = Hungary : <=50K (0/0) [0/0]  
 | | | | | | | native.country = Holand-Netherlands : <=50K (0/0) [0/0]  
 | | | capital.gain >= 5095.5 : >50K (49.26/1.13) [26/0]  
 | | | education.num >= 9 : >50K (256/-0) [134/0]  
 | | | education = 11th : <=50K (63.69/7) [27.52/5]  
 | | | education = Masters : >50K (26.21/7.17) [10.17/2.52]  
 | | | education = 9th : <=50K (40.78/2.13) [18.91/2]  
 | | | education = Some-college  
 | | | education.num < 10.03  
 | | | education.num < 9.97 : >50K (33/0) [19/0]  
 | | | education.num >= 9.97  
 | | | capital.gain < 5095.5  
 | | | fnlwgt < 212924  
 | | | native.country = United-States  
 | | | age < 33.5 : <=50K (57.39/6) [27/9]  
 | | | age >= 33.5  
 | | | workclass = Self-emp-not-inc : <=50K (25.33/4.99) [11.3/3.07]  
 | | | workclass = Private  
 | | | age < 48.27  
 | | | house.per.week < 43.45 : <=50K (57.15/14.93) [19.03/9]  
 | | | house.per.week >= 43.45



	fmlwgt	< 167243.5	
	age	< 35.5 : <=50K (6/0)	[3/1]
	age	>= 35.5	
	fmlwgt	< 61538.5 : <=50K (4/1)	[1/0]
	fmlwgt	>= 61538.5 : >50K (14/5)	[7.19/1.09]
	fmlwgt	>= 167243.5 : >50K (4/0)	[10/6]
	age	>= 48.27	
	house.per.week	< 51	
	house.per.week	< 27 : <=50K (2.13/0.09)	[0.47/0.09]
	house.per.week	>= 27 : >50K (34.5/9.38)	[16.75/5.56]
	house.per.week	>= 51 : <=50K (6/1)	[0/0]
	workclass	= State-gov : <=50K (2.98/1.01)	[2.04/1.01]
	workclass	= Federal-gov : <=50K (9.07/1.02)	[1.11/0.03]
	workclass	= Local-gov : <=50K (5.08/0.01)	[3.06/0.01]
	workclass	= Self-emp-inc	
	house.per.week	< 45 : <=50K (3.09/0.01)	[0.05/0.01]
	house.per.week	>= 45 : >50K (3/0)	[3.02/1.01]
	workclass	= Without-pay : <=50K (0/0)	[0/0]
	workclass	= Never-worked : <=50K (0/0)	[0/0]
	native.country	= Cuba : <=50K (0/0)	[0/0]
	native.country	= Jamaica : >50K (1.02/0.01)	[0.01/0.01]
	native.country	= India : <=50K (0/0)	[0/0]
	native.country	= Mexico : >50K (1.02/0.01)	[0.01/0.01]
	native.country	= South : >50K (0.13/0)	[0/0]
	native.country	= Puerto-Rico : <=50K (1.02/0.01)	[0.01/0]
	native.country	= Honduras : <=50K (0/0)	[0/0]
	native.country	= England : <=50K (0/0)	[0/0]
	native.country	= Canada : <=50K (2.05/1.02)	[0.02/0]
	native.country	= Germany : >50K (1.02/0.01)	[2.01/1.01]
	native.country	= Iran : >50K (1.02/0.01)	[0.01/0.01]
	native.country	= Philippines : >50K (1.16/0.01)	[1.01/0.01]
	native.country	= Italy : <=50K (1.02/0.01)	[1.01/0]
	native.country	= Poland : <=50K (2.05/0.02)	[0.02/0]
	native.country	= Columbia : <=50K (1.02/0.01)	[1.01/0]
	native.country	= Cambodia : <=50K (0/0)	[0/0]
	native.country	= Thailand : <=50K (0/0)	[0/0]
	native.country	= Ecuador : <=50K (0/0)	[0/0]
	native.country	= Laos : <=50K (0/0)	[0/0]
	native.country	= Taiwan : <=50K (0/0)	[0/0]
	native.country	= Haiti : <=50K (0/0)	[0/0]
	native.country	= Portugal : <=50K (0/0)	[0/0]

						native.country	=	Dominican-Republic : <=50K (0/0) [0/0]
						native.country	=	El-Salvador : <=50K (1.02/0.01) [0.01/0]
						native.country	=	France : <=50K (0/0) [0/0]
						native.country	=	Guatemala : <=50K (0/0) [0/0]
						native.country	=	China : <=50K (0/0) [0/0]
						native.country	=	Japan : <=50K (0/0) [0/0]
						native.country	=	Yugoslavia : <=50K (0/0) [0/0]
						native.country	=	Peru : <=50K (0/0) [0/0]
						native.country	=	Outlying-US(Guam-USVI-etc) : <=50K (0/0) [0/0]
						native.country	=	Scotland : <=50K (0/0) [0/0]
						native.country	=	Trinidad&Tobago : <=50K (0/0) [0/0]
						native.country	=	Greece : <=50K (1.02/0.01) [0.01/0]
						native.country	=	Nicaragua : <=50K (0/0) [0/0]
						native.country	=	Vietnam : <=50K (1.02/0.01) [0.01/0]
						native.country	=	Hong : <=50K (1.02/0.01) [0.01/0]
						native.country	=	Ireland : <=50K (0/0) [0/0]
						native.country	=	Hungary : <=50K (0/0) [0/0]
						native.country =		Holand-Netherlands : <=50K (0/0) [0/0]
						fnlwgt >= 212924		
						age < 40.07		
						age < 24.5 : <=50K (4/0) [3.13/0]		
						age >= 24.5		
						race = White		
						capital.gain < 2966		
						age < 39		
						age < 37.49		
						house.per.week < 47.5 : <=50K (26.13/10) [21/3]		
						house.per.week >= 47.5		
						house.per.week < 65 : >50K (10/2) [5/2]		
						house.per.week >= 65 : <=50K (2/0) [1/0]		
						age >= 37.49 : >50K (2/0) [1/0]		
						age >= 39 : <=50K (2/0) [1/1]		
						capital.gain >= 2966 : >50K (3/0) [1/1]		
						race = Black : <=50K (3/1) [2/1]		
						race = Asian-Pac-Islander : <=50K (2/0) [1/1]		
						race = Amer-Indian-Eskimo : <=50K (0/0) [0/0]		
						race = Other : >50K (1/0) [0/0]		
						age >= 40.07 : >50K (60.95/14.17) [37.65/13.26]		
						capital.gain >= 5095.5 : >50K (26.39/1) [7.26/0]		
						education.num >= 10.03 : >50K (67/0) [35/0]		
						education = Assoc-acdm		

```

| | | age < 37.5 : <=50K (18.26/3) [9.13/2]
| | | age >= 37.5
| | | | workclass = Self-emp-not-inc : <=50K (3.06/0.03) [1.05/1.03]
| | | | workclass = Private
| | | | | fnlwgt < 153534.5 : <=50K (8.09/3.09) [3.09/0.09]
| | | | | fnlwgt >= 153534.5 : >50K (10.28/1.19) [3.19/1.09]
| | | | workclass = State-gov : >50K (0/0) [0/0]
| | | | workclass = Federal-gov : >50K (1.02/0.01) [1.02/0.01]
| | | | workclass = Local-gov : >50K (1.02/0.01) [0.02/0.01]
| | | | workclass = Self-emp-inc : <=50K (2.04/0.02) [1.03/1.02]
| | | | workclass = Without-pay : >50K (0/0) [0/0]
| | | | workclass = Never-worked : >50K (0/0) [0/0]
| | | education = Assoc-voc
| | | age < 34 : <=50K (32.26/8.13) [15/5]
| | | age >= 34
| | | | native.country = United-States
| | | | | capital.loss < 943.5
| | | | | | fnlwgt < 182905
| | | | | | | capital.gain < 6155.5 : <=50K (43.45/18.33) [15/3]
| | | | | | | capital.gain >= 6155.5 : >50K (3/-0) [4/0]
| | | | | | | fnlwgt >= 182905
| | | | | | | workclass = Self-emp-not-inc : >50K (0/0) [0/0]
| | | | | | | workclass = Private
| | | | | | | | fnlwgt < 238322 : >50K (9.94/0) [7/5]
| | | | | | | | fnlwgt >= 238322 : >50K (18.13/7) [8.1/2.1]
| | | | | | | | workclass = State-gov : >50K (1.01/0) [1/1]
| | | | | | | | workclass = Federal-gov : >50K (4.03/2) [0.01/0.01]
| | | | | | | | workclass = Local-gov : >50K (3.02/0) [1.01/1.01]
| | | | | | | | workclass = Self-emp-inc : <=50K (1.01/0.01) [0/0]
| | | | | | | | workclass = Without-pay : >50K (0/0) [0/0]
| | | | | | | | workclass = Never-worked : >50K (0/0) [0/0]
| | | | | capital.loss >= 943.5 : >50K (6.13/-0) [1/1]
| | | | native.country = Cuba : >50K (0/0) [0/0]
| | | | native.country = Jamaica : >50K (0/0) [0/0]
| | | | native.country = India : >50K (0/0) [0/0]
| | | | native.country = Mexico : >50K (0/0) [0/0]
| | | | native.country = South : >50K (0/0) [0/0]
| | | | native.country = Puerto-Rico : >50K (0/0) [0/0]
| | | | native.country = Honduras : >50K (0/0) [0/0]
| | | | native.country = England : >50K (0/0) [0/0]
| | | | native.country = Canada : >50K (2.06/0) [1.13/1.13]

```

| | | native.country = Germany : >50K (0/0) [0/0]  
 | | | native.country = Iran : >50K (0/0) [0/0]  
 | | | native.country = Philippines : >50K (0/0) [0/0]  
 | | | native.country = Italy : >50K (0/0) [0/0]  
 | | | native.country = Poland : >50K (0/0) [1/0]  
 | | | native.country = Columbia : >50K (0/0) [0/0]  
 | | | native.country = Cambodia : >50K (0/0) [0/0]  
 | | | native.country = Thailand : >50K (0/0) [0/0]  
 | | | native.country = Ecuador : >50K (0/0) [0/0]  
 | | | native.country = Laos : >50K (0/0) [0/0]  
 | | | native.country = Taiwan : >50K (0/0) [0/0]  
 | | | native.country = Haiti : >50K (0/0) [0/0]  
 | | | native.country = Portugal : >50K (1.03/0) [0/0]  
 | | | native.country = Dominican-Republic : >50K (0/0) [0/0]  
 | | | native.country = El-Salvador : >50K (0/0) [0/0]  
 | | | native.country = France : >50K (0/0) [0/0]  
 | | | native.country = Guatemala : >50K (0/0) [0/0]  
 | | | native.country = China : >50K (0/0) [0/0]  
 | | | native.country = Japan : >50K (0/0) [0/0]  
 | | | native.country = Yugoslavia : <=50K (1.03/0.03) [0/0]  
 | | | native.country = Peru : >50K (0/0) [0/0]  
 | | | native.country = Outlying-US(Guam-USVI-etc) : >50K (0/0) [0/0]  
 | | | native.country = Scotland : >50K (0/0) [0/0]  
 | | | native.country = Trinidad&Tobago : >50K (0/0) [0/0]  
 | | | native.country = Greece : <=50K (1.03/0.03) [0/0]  
 | | | native.country = Nicaragua : >50K (0/0) [0/0]  
 | | | native.country = Vietnam : <=50K (1.03/0.03) [0/0]  
 | | | native.country = Hong : >50K (0/0) [0/0]  
 | | | native.country = Ireland : >50K (0/0) [0/0]  
 | | | native.country = Hungary : >50K (0/0) [1/0]  
 | | | native.country = Holand-Netherlands : >50K (0/0) [0/0]  
 | | education = 7th-8th : <=50K (58.99/8.13) [31.78/0.13]  
 | | education = Doctorate : >50K (2.04/1.13) [0.52/0.13]  
 | | education = Prof-school : >50K (7.04/2.52) [0.52/0.26]  
 | | education = 5th-6th : <=50K (15.43/1.13) [10.78/2.13]  
 | | education = 10th : <=50K (72.3/12.26) [27.39/5]  
 | | education = 1st-4th : <=50K (9.52/2) [3.13/0]  
 | | education = Preschool : <=50K (3.26/0) [0/0]  
 | | education = 12th : <=50K (17.52/5) [18/3]  
 | occupation = Transport-moving  
 | | capital.gain < 3032

```

|| | | education = Bachelors
|| | | | education.num < 12.72 : >50K (4/0) [4/0]
|| | | | education.num >= 12.72 : <=50K (26.37/12.46) [10.88/3.23]
|| | | education = HS-grad
|| | | | education.num < 9
|| | | | | age < 45.02
|| | | | | | race = White
|| | | | | | | house.per.week < 71.5
|| | | | | | | | workclass = Self-emp-not-inc : <=50K (14.06/3) [8.02/1]
|| | | | | | | | workclass = Private
|| | | | | | | | | fnlwgt < 213860
|| | | | | | | | | | fnlwgt < 201889
|| | | | | | | | | | | fnlwgt < 97029.5 : <=50K (20.03/8) [8.06/2]
|| | | | | | | | | | | fnlwgt >= 97029.5
|| | | | | | | | | | | | house.per.week < 42.5 : <=50K (35.26/2) [11.03/2]
|| | | | | | | | | | | | house.per.week >= 42.5
|| | | | | | | | | | | | | fnlwgt < 129570.5 : <=50K (9/0) [2/1]
|| | | | | | | | | | | | | fnlwgt >= 129570.5
|| | | | | | | | | | | | | | age < 31.5 : <=50K (11/1) [2/0]
|| | | | | | | | | | | | | | age >= 31.5 : >50K (17.03/8.03) [13/4]
|| | | | | | | | | | | | | | fnlwgt >= 201889 : <=50K (15.03/7) [3/1]
|| | | | | | | | | | | | | | fnlwgt >= 213860 : <=50K (42.26/4.03) [21.1/7]
|| | | | | | | | | | | | | | workclass = State-gov : >50K (2.01/0.01) [1/0]
|| | | | | | | | | | | | | | workclass = Federal-gov : <=50K (3.01/0) [1/1]
|| | | | | | | | | | | | | | workclass = Local-gov : <=50K (6.02/1) [6.01/0]
|| | | | | | | | | | | | | | workclass = Self-emp-inc : <=50K (2.01/0) [0/0]
|| | | | | | | | | | | | | | workclass = Without-pay : <=50K (0/0) [0/0]
|| | | | | | | | | | | | | | workclass = Never-worked : <=50K (0/0) [0/0]
|| | | | | | | | | | | | | | house.per.week >= 71.5 : <=50K (3.04/3) [7/2]
|| | | | | | | | | | | | | | race = Black : <=50K (15.08/2) [10.15/2]
|| | | | | | | | | | | | | | race = Asian-Pac-Islander : <=50K (4.04/0) [1/1]
|| | | | | | | | | | | | | | race = Amer-Indian-Eskimo : <=50K (7/0) [1/0]
|| | | | | | | | | | | | | | race = Other : <=50K (3/0) [1/0]
|| | | | | | | | | | | | | | age >= 45.02
|| | | | | | | | | | | | | | fnlwgt < 74469.5 : <=50K (16.42/2.15) [9.31/4.08]
|| | | | | | | | | | | | | | fnlwgt >= 74469.5 : <=50K (100.79/48.53) [64.79/21.34]
|| | | | | | | | | | | | | | education.num >= 9 : >50K (28/0) [16/0]
|| | | | | | | | | | | | | | education = 11th : <=50K (26.5/0) [21.15/3]
|| | | | | | | | | | | | | | education = Masters : >50K (4.61/2.34) [3.31/1.15]
|| | | | | | | | | | | | | | education = 9th : <=50K (16.15/2) [6.27/0]
|| | | | | | | | | | | | | | education = Some-college

```

```

|| | | workclass = Self-emp-not-inc : <=50K (7.1/2.03) [3.05/0.02]
|| | | workclass = Private
|| | | | age < 47.5
|| | | | | capital.loss < 952
|| | | | | | fnlwgt < 398335.5 : <=50K (59.22/14.03) [23.15/6.03]
|| | | | | | fnlwgt >= 398335.5 : >50K (4.03/1.03) [1/0]
|| | | | | capital.loss >= 952 : >50K (2/0) [4/2]
|| | | | | age >= 47.5
|| | | | | | fnlwgt < 156797 : <=50K (7.37/2.06) [2.28/0.03]
|| | | | | | fnlwgt >= 156797
|| | | | | | | age < 54.5 : >50K (8.15/0.06) [2/0]
|| | | | | | | age >= 54.5
|| | | | | | | | fnlwgt < 185179.24 : >50K (3.06/0.06) [1.06/0]
|| | | | | | | | fnlwgt >= 185179.24 : <=50K (6.4/2.15) [2.19/0.09]
|| | | | workclass = State-gov : <=50K (1.01/0) [2.01/0]
|| | | | workclass = Federal-gov : >50K (2.03/0.02) [1.02/1.01]
|| | | | workclass = Local-gov : <=50K (7.1/1.03) [4.05/0.02]
|| | | | workclass = Self-emp-inc : >50K (4.06/1.04) [1.03/0.02]
|| | | | workclass = Without-pay : <=50K (0/0) [0/0]
|| | | | workclass = Never-worked : <=50K (0/0) [0/0]
|| | | education = Assoc-acdm : <=50K (7.19/0.04) [4.15/1.08]
|| | | education = Assoc-voc : <=50K (13.27/2.23) [8.08/2]
|| | | education = 7th-8th : <=50K (30.88/5.04) [16.15/1]
|| | | education = Doctorate : <=50K (1.23/0.19) [0.15/0.11]
|| | | education = Prof-school : >50K (1.23/0.15) [0.15/0.08]
|| | | education = 5th-6th : <=50K (13.42/2.04) [4.19/0]
|| | | education = 10th : <=50K (35.34/11.04) [16.11/3]
|| | | education = 1st-4th : <=50K (4.11/0) [3.04/0]
|| | | education = Preschool : <=50K (0.08/0) [0/0]
|| | | education = 12th : <=50K (10.15/2) [4/1]
|| | capital.gain >= 3032
|| | | capital.gain < 5039.36
|| | | | age < 40 : <=50K (7.08/1.04) [2/0]
|| | | | age >= 40 : >50K (21.31/3.27) [6.15/4.11]
|| | | capital.gain >= 5039.36 : >50K (38.76/1.08) [22.27/0]
| occupation = Farming-fishing
|| | capital.gain < 5095.5
|| | | education = Bachelors
|| | | | fnlwgt < 34586.5 : >50K (3.04/0.04) [3/0]
|| | | | fnlwgt >= 34586.5 : <=50K (25.7/4.25) [16.5/6.14]
|| | | education = HS-grad

```

```

|| | | | fnlwgt < 38558
|| | | | fnlwgt < 33225 : <=50K (18.12/2.04) [4.04/1.02]
|| | | | fnlwgt >= 33225
|| | | | | fnlwgt < 34096.5 : >50K (8/0) [1/0]
|| | | | | fnlwgt >= 34096.5 : <=50K (13.02/6) [3.02/1.02]
|| | | | fnlwgt >= 38558
|| | | | | house.per.week < 52.5 : <=50K (90.98/2.35) [33.24/3.14]
|| | | | | house.per.week >= 52.5 : <=50K (33.17/7.02) [26.06/8.04]
|| | | education = 11th : <=50K (10.27/0) [7.08/2]
|| | | education = Masters : >50K (11.33/4.19) [5.17/0.08]
|| | | education = 9th : <=50K (11.1/0) [4.14/1]
|| | | education = Some-college : <=50K (56.87/8.23) [36.47/7.14]
|| | | education = Assoc-acdm : <=50K (7.1/2.02) [1.08/0.04]
|| | | education = Assoc-voc : <=50K (21.19/5.14) [15.04/4]
|| | | education = 7th-8th : <=50K (27.47/2.02) [15.1/0]
|| | | education = Doctorate : >50K (0.12/0.02) [0.08/0.02]
|| | | education = Prof-school : <=50K (2.12/1.04) [1.08/0.04]
|| | | education = 5th-6th : <=50K (12.23/0.02) [5.1/0]
|| | | education = 10th : <=50K (11.19/0.02) [6.06/0]
|| | | education = 1st-4th : <=50K (7.08/1) [1.02/0]
|| | | education = Preschool : <=50K (2.04/0) [0/0]
|| | | education = 12th : <=50K (2.08/0) [2/0]
|| | capital.gain >= 5095.5 : >50K (19.41/2.04) [8.14/0]
| occupation = Machine-op-inspct
|| | education = Bachelors
|| | | education.num < 12.65 : >50K (10/0) [6/0]
|| | | education.num >= 12.65
|| | | | race = White
|| | | | | fnlwgt < 176768
|| | | | | fnlwgt < 35449 : >50K (2.1/0.07) [0/0]
|| | | | | fnlwgt >= 35449
|| | | | | | fnlwgt < 110265 : <=50K (5.14/0.03) [2.21/0]
|| | | | | | fnlwgt >= 110265
|| | | | | | | fnlwgt < 161257 : >50K (2.28/0.17) [8.24/3.17]
|| | | | | | | fnlwgt >= 161257 : <=50K (2.17/0) [0.03/0]
|| | | | | fnlwgt >= 176768
|| | | | | fnlwgt < 183775 : >50K (4.03/0) [3.14/0.07]
|| | | | | fnlwgt >= 183775 : >50K (7.63/2.24) [5.24/1.07]
|| | | | race = Black : <=50K (2.07/0.03) [0/0]
|| | | | race = Asian-Pac-Islander : >50K (1.07/0.07) [0.03/0.03]
|| | | | race = Amer-Indian-Eskimo : <=50K (1/0) [0/0]

```

```

| | | | race = Other : <=50K (2/0) [0/0]
| | education = HS-grad
| | | education.num < 9.03
| | | | capital.gain < 4102
| | | | | age < 28.5 : <=50K (44.07/1) [21/1]
| | | | | age >= 28.5
| | | | | | house.per.week < 40.82 : <=50K (197.25/43.49) [93.03/26.28]
| | | | | | house.per.week >= 40.82
| | | | | | | education.num < 8.7 : >50K (4/0) [3/0]
| | | | | | | education.num >= 8.7
| | | | | | | | capital.loss < 1592.66 : <=50K (55.56/24.21) [45.28/22.1]
| | | | | | | | capital.loss >= 1592.66 : >50K (4/0) [1/0]
| | | | capital.gain >= 4102 : >50K (12.07/1.03) [10/3]
| | | education.num >= 9.03 : >50K (35/0) [19/0]
| | education = 11th : <=50K (27.45/2) [13.14/1]
| | education = Masters : <=50K (2.59/1.28) [3.31/0.17]
| | education = 9th : <=50K (20.21/1.03) [13.24/1]
| | education = Some-college
| | | capital.gain < 5211 : <=50K (97.47/23.38) [36.8/11.24]
| | | capital.gain >= 5211 : >50K (4.1/0) [1.07/0]
| | education = Assoc-acdm
| | | age < 45.5 : <=50K (6.07/1) [3.03/1]
| | | age >= 45.5 : >50K (4.14/0.07) [3.1/1.03]
| | education = Assoc-voc : <=50K (21.31/11.24) [13.07/3]
| | education = 7th-8th : <=50K (25.8/2.03) [13.21/3.03]
| | education = Doctorate : >50K (0.28/0.03) [0.14/0.03]
| | education = Prof-school : <=50K (1.28/0.14) [0.14/0.07]
| | education = 5th-6th : <=50K (17.38/1.03) [5.21/0.03]
| | education = 10th : <=50K (31.35/5.07) [12.1/1]
| | education = 1st-4th : <=50K (7.14/1) [4.03/0]
| | education = Preschool : <=50K (2.07/0) [0/0]
| | education = 12th : <=50K (6.14/1) [3/0]
| occupation = Tech-support
| | fnlwgt < 108785.5
| | | age < 44
| | | | workclass = Self-emp-not-inc : <=50K (2.01/0) [0.01/0]
| | | | workclass = Private
| | | | | fnlwgt < 65224.5 : <=50K (6.06/0.01) [7.04/2]
| | | | | fnlwgt >= 65224.5
| | | | | | education = Bachelors : >50K (3/2) [3/0]
| | | | | | education = HS-grad : <=50K (6.02/2) [0/0]

```



|| | | | | education = 11th : <=50K (0.01/0) [0/0]  
 || | | | | education = Masters : <=50K (0/0) [0/0]  
 || | | | | education = 9th : <=50K (0/0) [0/0]  
 || | | | | education = Some-college : <=50K (2.01/1) [2.01/0]  
 || | | | | education = Assoc-acdm : <=50K (2/0) [1/0]  
 || | | | | education = Assoc-voc : >50K (1/0) [1/0]  
 || | | | | education = 7th-8th : <=50K (0/0) [0/0]  
 || | | | | education = Doctorate : <=50K (0/0) [0/0]  
 || | | | | education = Prof-school : <=50K (0/0) [0/0]  
 || | | | | education = 5th-6th : <=50K (0/0) [0/0]  
 || | | | | education = 10th : <=50K (0/0) [0/0]  
 || | | | | education = 1st-4th : <=50K (0/0) [0/0]  
 || | | | | education = Preschool : <=50K (0/0) [0/0]  
 || | | | | education = 12th : <=50K (0.01/0) [0/0]  
 || | | | workclass = State-gov : <=50K (5.03/0) [0.01/0]  
 || | | | workclass = Federal-gov : <=50K (2.01/1) [2.01/1]  
 || | | | workclass = Local-gov : <=50K (2.01/1) [0.01/0]  
 || | | | workclass = Self-emp-inc : <=50K (1.01/0) [0/0]  
 || | | | workclass = Without-pay : <=50K (0/0) [0/0]  
 || | | | workclass = Never-worked : <=50K (0/0) [0/0]  
 || | | age >= 44 : >50K (27.99/10.69) [12.52/4.37]  
 || | fnlwgt >= 108785.5  
 || | | education = Bachelors : >50K (89.61/12.34) [47.34/7.19]  
 || | | education = HS-grad : >50K (40.53/13.3) [18.89/6.76]  
 || | | education = 11th : <=50K (1.15/0) [0.03/0]  
 || | | education = Masters : >50K (17.22/3.12) [2.13/0.07]  
 || | | education = 9th : <=50K (1.07/0.02) [0.1/0]  
 || | | education = Some-college : >50K (64.67/15.44) [34.29/9.17]  
 || | | education = Assoc-acdm  
 || | | | age < 40.5 : >50K (10.02/1.02) [6.02/3.02]  
 || | | | age >= 40.5  
 || | | | | fnlwgt < 207657.5 : >50K (3.02/1.02) [2.03/0]  
 || | | | | fnlwgt >= 207657.5 : <=50K (2.03/0.02) [1.02/0]  
 || | | | education = Assoc-voc  
 || | | | | age < 40.5  
 || | | | | | house.per.week < 41 : <=50K (8.02/2) [3.02/0]  
 || | | | | | house.per.week >= 41 : >50K (4.03/0) [1/1]  
 || | | | | age >= 40.5 : >50K (13.03/0) [10.02/2.02]  
 || | | | education = 7th-8th : <=50K (1.2/0) [1.08/0.02]  
 || | | | education = Doctorate : >50K (2.1/1.02) [0.03/0.02]  
 || | | | education = Prof-school : <=50K (2.12/1.05) [0.05/0.02]

```

| | | education = 5th-6th : <=50K (0.15/0.02) [0.1/0.02]
| | | education = 10th : >50K (0.15/0.12) [1.03/0.03]
| | | education = 1st-4th : <=50K (0.05/0) [0.02/0]
| | | education = Preschool : <=50K (0.03/0) [0/0]
| | | education = 12th : <=50K (1.03/0) [0/0]
| | occupation = Protective-serv
| | | education = Bachelors : >50K (56.65/13.38) [23.39/4.26]
| | | education = HS-grad
| | | | capital.gain < 7032.5
| | | | | fnlwgt < 139619 : <=50K (28.6/4.15) [9.3/3.06]
| | | | | fnlwgt >= 139619
| | | | | | fnlwgt < 149022 : >50K (3.06/0.05) [2.02/1]
| | | | | | fnlwgt >= 149022 : <=50K (49.04/18.15) [27.68/11.09]
| | | | capital.gain >= 7032.5 : >50K (3/0) [2/0]
| | | education = 11th : <=50K (1.2/0) [1.06/1]
| | | education = Masters : >50K (3.26/0.14) [3.14/1.06]
| | | education = 9th : <=50K (2.09/0.02) [0.11/0]
| | | education = Some-college
| | | | education.num < 10.39
| | | | | age < 35.5 : <=50K (28.08/7.02) [13.03/5]
| | | | | age >= 35.5
| | | | | | age < 58.5
| | | | | | | workclass = Self-emp-not-inc : >50K (2.01/0.01) [0/0]
| | | | | | | workclass = Private
| | | | | | | | race = White
| | | | | | | | | age < 44.1 : <=50K (2.01/1) [0.01/0]
| | | | | | | | | age >= 44.1 : >50K (4.03/0.02) [4.01/1.01]
| | | | | | | | | race = Black : >50K (4/0) [0/0]
| | | | | | | | | race = Asian-Pac-Islander : >50K (0/0) [0/0]
| | | | | | | | | race = Amer-Indian-Eskimo : >50K (0/0) [0/0]
| | | | | | | | | race = Other : <=50K (1/0) [0/0]
| | | | | | | workclass = State-gov
| | | | | | | | house.per.week < 45 : >50K (5.03/0.02) [2.01/2]
| | | | | | | | house.per.week >= 45 : <=50K (3.01/1) [1.01/0]
| | | | | | | | workclass = Federal-gov : >50K (2.01/0.01) [1/1]
| | | | | | | | workclass = Local-gov : >50K (27.11/8.08) [12.06/2.03]
| | | | | | | | workclass = Self-emp-inc : >50K (1/0) [0/0]
| | | | | | | | workclass = Without-pay : >50K (0/0) [0/0]
| | | | | | | | workclass = Never-worked : >50K (0/0) [0/0]
| | | | | | | | age >= 58.5 : <=50K (10.39/1.14) [3.24/2.09]
| | | | education.num >= 10.39 : >50K (12/0) [1/0]

```

- | | education = Assoc-acdm
- | | | workclass = Self-emp-not-inc : <=50K (1.01/0) [0/0]
- | | | workclass = Private : <=50K (2.01/1) [1.01/0]
- | | | workclass = State-gov : <=50K (2.01/0) [2.01/0]
- | | | workclass = Federal-gov : <=50K (1.01/0) [0/0]
- | | | workclass = Local-gov : >50K (9.05/3.04) [5.04/1.02]
- | | | workclass = Self-emp-inc : <=50K (0/0) [0/0]
- | | | workclass = Without-pay : <=50K (0/0) [0/0]
- | | | workclass = Never-worked : <=50K (0/0) [0/0]
- | | education = Assoc-voc : >50K (18.14/8.03) [17.03/5.03]
- | | education = 7th-8th : <=50K (3.35/0.02) [4.09/0.02]
- | | education = Doctorate : >50K (0.12/0.02) [0.06/0.02]
- | | education = Prof-school : >50K (0.12/0.06) [1.06/0.03]
- | | education = 5th-6th : <=50K (0.17/0.02) [1.09/0.02]
- | | education = 10th : <=50K (1.15/0.03) [1.05/0]
- | | education = 1st-4th : <=50K (0.06/0) [0.02/0]
- | | education = Preschool : <=50K (0.03/0) [0/0]
- | | education = 12th : <=50K (1.06/0) [1/0]
- | occupation = Armed-Forces : <=50K (2.03/1.01) [0.02/0]
- | occupation = Priv-house-serv : >50K (0/0) [0/0]
- relationship = Not-in-family
- | capital.gain < 7599.39
- | | occupation = Exec-managerial
- | | | marital.status = Married-civ-spouse : >50K (47.13/0.13) [20.26/0.26]
- | | | marital.status = Divorced
- | | | education = Bachelors
- | | | | workclass = Self-emp-not-inc : <=50K (10.27/0) [1/0]
- | | | | workclass = Private
- | | | | race = White
- | | | | | fnlwgt < 106369 : <=50K (3.08/0) [2/0]
- | | | | | fnlwgt >= 106369
- | | | | | house.per.week < 42.5 : <=50K (8.48/2) [10/2]
- | | | | | house.per.week >= 42.5 : >50K (21.4/6.4) [11/4]
- | | | | race = Black : >50K (2/0) [0/0]
- | | | | race = Asian-Pac-Islander : <=50K (0/0) [1/0]
- | | | | race = Amer-Indian-Eskimo : <=50K (1/0) [0/0]
- | | | | race = Other : <=50K (0/0) [0/0]
- | | | | workclass = State-gov : <=50K (2.05/0) [3/1]
- | | | | workclass = Federal-gov : <=50K (4.11/1) [1/1]
- | | | | workclass = Local-gov : <=50K (2.05/1) [0/0]
- | | | | workclass = Self-emp-inc : <=50K (4.11/1) [1/0]

| | | | workclass = Without-pay : <=50K (0/0) [0/0]  
 | | | | workclass = Never-worked : <=50K (0/0) [0/0]  
 | | | | education = HS-grad  
 | | | | fnlwgt < 336147.5 : <=50K (54.72/2.26) [25.29/4]  
 | | | | fnlwgt >= 336147.5 : >50K (7.13/3.13) [1/0]  
 | | | | education = 11th : <=50K (1.52/0) [1.26/0]  
 | | | | education = Masters  
 | | | | workclass = Self-emp-not-inc : <=50K (2.02/1) [0/0]  
 | | | | workclass = Private : >50K (6.05/1.05) [9/3]  
 | | | | workclass = State-gov : <=50K (2.02/0) [0/0]  
 | | | | workclass = Federal-gov : >50K (1.01/0.01) [0/0]  
 | | | | workclass = Local-gov : <=50K (4.03/1) [1/0]  
 | | | | workclass = Self-emp-inc : <=50K (1.01/1) [3/1]  
 | | | | workclass = Without-pay : >50K (0/0) [0/0]  
 | | | | workclass = Never-worked : >50K (0/0) [0/0]  
 | | | | education = 9th : <=50K (0.78/0) [0.26/0]  
 | | | | education = Some-college : <=50K (56.42/6.13) [36.16/5]  
 | | | | education = Assoc-acdm : <=50K (10.39/1) [8/2]  
 | | | | education = Assoc-voc : <=50K (9.13/1) [6.39/1]  
 | | | | education = 7th-8th : <=50K (0.78/0) [2.13/0]  
 | | | | education = Doctorate : >50K (3/1) [1/0]  
 | | | | education = Prof-school : <=50K (3.13/1) [2.13/1]  
 | | | | education = 5th-6th : <=50K (0/0) [0.13/0]  
 | | | | education = 10th : <=50K (1.91/0) [0.65/0]  
 | | | | education = 1st-4th : <=50K (0.26/0) [0/0]  
 | | | | education = Preschool : <=50K (0/0) [0/0]  
 | | | | education = 12th : <=50K (1.13/0) [0.13/0]  
 | | | marital.status = Married-spouse-absent : <=50K (6.42/1.13) [13.78/5]  
 | | | marital.status = Never-married  
 | | | | education = Bachelors : <=50K (152.33/21.13) [63.03/14]  
 | | | | education = HS-grad : <=50K (58.49/2) [20.42/1]  
 | | | | education = 11th : <=50K (2.78/0) [0.65/0]  
 | | | | education = Masters : <=50K (45.52/13.13) [15.52/3]  
 | | | | education = 9th : <=50K (1.26/0) [0.13/0]  
 | | | | education = Some-college : <=50K (67.79/4) [35.16/3]  
 | | | | education = Assoc-acdm : <=50K (9.65/1) [5.39/1]  
 | | | | education = Assoc-voc : <=50K (4.39/1) [5/1]  
 | | | | education = 7th-8th : <=50K (0.52/0) [0.13/0]  
 | | | | education = Doctorate : >50K (3.13/0.13) [3.13/0]  
 | | | | education = Prof-school : >50K (5.13/2) [1/0]  
 | | | | education = 5th-6th : <=50K (0.13/0) [0/0]

| | | education = 10th : <=50K (0.52/0) [0/0]  
 | | | education = 1st-4th : <=50K (1/0) [0.13/0]  
 | | | education = Preschool : <=50K (0/0) [0/0]  
 | | | education = 12th : <=50K (4.65/1) [0.26/0]  
 | | marital.status = Separated : <=50K (27.33/5) [15.03/3]  
 | | marital.status = Married-AF-spouse : <=50K (0/0) [0/0]  
 | | marital.status = Widowed : <=50K (45.57/5.26) [23.66/4]  
 | | occupation = Handlers-cleaners : <=50K (226.35/4.37) [100.49/1.04]  
 | | occupation = Prof-specialty  
 | | age < 29.5 : <=50K (279.92/6) [124/10]  
 | | age >= 29.5  
 | | | education.num < 14.05  
 | | | | marital.status = Married-civ-spouse : >50K (11/1) [6.15/0.15]  
 | | | | marital.status = Divorced  
 | | | | | house.per.week < 44.5 : <=50K (126.08/13.31) [49.69/6]  
 | | | | | house.per.week >= 44.5  
 | | | | | workclass = Self-emp-not-inc : <=50K (8.41/3.03) [2.07/0]  
 | | | | | workclass = Private : <=50K (14.72/7.06) [19.12/3]  
 | | | | | workclass = State-gov : <=50K (3.15/0.01) [0.03/0]  
 | | | | | workclass = Federal-gov : >50K (1.05/0.05) [1.01/0.01]  
 | | | | | workclass = Local-gov : <=50K (8.41/0.03) [6.07/1]  
 | | | | | workclass = Self-emp-inc : >50K (2.1/1.09) [1.02/0.02]  
 | | | | | workclass = Without-pay : <=50K (0/0) [0/0]  
 | | | | | workclass = Never-worked : <=50K (0/0) [0/0]  
 | | | | marital.status = Married-spouse-absent : <=50K (16.23/2.15) [10.46/3]  
 | | | | marital.status = Never-married  
 | | | | | native.country = United-States : <=50K (222.45/42.23) [147.76/28.85]  
 | | | | | native.country = Cuba : <=50K (0/0) [2/0]  
 | | | | | native.country = Jamaica : <=50K (0/0) [0/0]  
 | | | | | native.country = India : <=50K (0/0) [0/0]  
 | | | | | native.country = Mexico : <=50K (7.18/0.03) [0.06/0.06]  
 | | | | | native.country = South : <=50K (0/0) [0/0]  
 | | | | | native.country = Puerto-Rico : <=50K (0/0) [0/0]  
 | | | | | native.country = Honduras : <=50K (0/0) [0/0]  
 | | | | | native.country = England : <=50K (1.03/0) [0.01/0.01]  
 | | | | | native.country = Canada : <=50K (2.05/0.01) [0.02/0.02]  
 | | | | | native.country = Germany : <=50K (3.08/0.01) [1.03/0.03]  
 | | | | | native.country = Iran : <=50K (0/0) [1/0]  
 | | | | | native.country = Philippines : >50K (1.03/0.02) [0.01/0]  
 | | | | | native.country = Italy : <=50K (0/0) [0/0]  
 | | | | | native.country = Poland : <=50K (0/0) [0/0]

| | | | | native.country = Columbia : <=50K (1.03/0) [0.01/0.01]  
 | | | | | native.country = Cambodia : <=50K (0/0) [0/0]  
 | | | | | native.country = Thailand : <=50K (0/0) [0/0]  
 | | | | | native.country = Ecuador : <=50K (0/0) [0/0]  
 | | | | | native.country = Laos : <=50K (0/0) [0/0]  
 | | | | | native.country = Taiwan : >50K (1.18/0.18) [0.01/0]  
 | | | | | native.country = Haiti : <=50K (0/0) [0/0]  
 | | | | | native.country = Portugal : <=50K (0/0) [0/0]  
 | | | | | native.country = Dominican-Republic : <=50K (0/0) [0/0]  
 | | | | | native.country = El-Salvador : <=50K (0/0) [0/0]  
 | | | | | native.country = France : <=50K (0/0) [1/0]  
 | | | | | native.country = Guatemala : <=50K (0/0) [0/0]  
 | | | | | native.country = China : <=50K (0/0) [0/0]  
 | | | | | native.country = Japan : <=50K (0/0) [1/0]  
 | | | | | native.country = Yugoslavia : <=50K (0/0) [0/0]  
 | | | | | native.country = Peru : <=50K (0/0) [0/0]  
 | | | | | native.country = Outlying-US(Guam-USVI-etc) : <=50K (0/0) [1/0]  
 | | | | | native.country = Scotland : <=50K (0/0) [0/0]  
 | | | | | native.country = Trinidad&Tobago : <=50K (1.03/0) [0.01/0.01]  
 | | | | | native.country = Greece : <=50K (0/0) [0/0]  
 | | | | | native.country = Nicaragua : <=50K (0/0) [0/0]  
 | | | | | native.country = Vietnam : <=50K (0/0) [0/0]  
 | | | | | native.country = Hong : <=50K (0/0) [0/0]  
 | | | | | native.country = Ireland : <=50K (1.03/0) [0.01/0.01]  
 | | | | | native.country = Hungary : <=50K (0/0) [0/0]  
 | | | | | native.country = Holand-Netherlands : <=50K (0/0) [0/0]  
 | | | | | marital.status = Separated : <=50K (20/5) [11.08/2]  
 | | | | | marital.status = Married-AF-spouse : <=50K (0/0) [0/0]  
 | | | | | marital.status = Widowed : <=50K (32.39/1.31) [24.39/3]  
 | | | | | education.num >= 14.05  
 | | | | | age < 30.5 : <=50K (4/0) [2/0]  
 | | | | | age >= 30.5 : >50K (74.46/31.31) [45.46/21.31]  
 | | | | | occupation = Other-service : <=50K (665.31/7.09) [345.15/8.12]  
 | | | | | occupation = Adm-clerical : <=50K (816.56/23.34) [393.82/8.15]  
 | | | | | occupation = Sales  
 | | | | | education = Bachelors  
 | | | | | age < 28.5 : <=50K (52.67/1) [35.67/1]  
 | | | | | age >= 28.5  
 | | | | | house.per.week < 46.71 : <=50K (54/9.22) [24.33/5]  
 | | | | | house.per.week >= 46.71  
 | | | | | marital.status = Married-civ-spouse : >50K (2/0) [0/0]

| | | | | marital.status = Divorced  
 | | | | | sex = Male : >50K (12.22/5.22) [3/1]  
 | | | | | sex = Female : <=50K (5.33/0) [4/1]  
 | | | | | marital.status = Married-spouse-absent : <=50K (0/0) [1.11/0]  
 | | | | | marital.status = Never-married : <=50K (26.11/10) [9.11/3]  
 | | | | | marital.status = Separated : >50K (1/0) [2/1]  
 | | | | | marital.status = Married-AF-spouse : <=50K (0/0) [0/0]  
 | | | | | marital.status = Widowed : >50K (1/0) [0.11/0.11]  
 | | education = HS-grad : <=50K (185.45/7.44) [88.67/6]  
 | | education = 11th : <=50K (14.45/0) [4.11/1]  
 | | education = Masters : <=50K (24.78/7.11) [10.56/2]  
 | | education = 9th : <=50K (6.22/0) [5.44/1]  
 | | education = Some-college  
 | | | marital.status = Married-civ-spouse : >50K (6/0) [2/0]  
 | | | marital.status = Divorced : <=50K (30.22/1.11) [15/1]  
 | | | marital.status = Married-spouse-absent : <=50K (1.11/0) [2.11/0]  
 | | | marital.status = Never-married : <=50K (97.11/1) [48/2]  
 | | | marital.status = Separated : <=50K (10.33/1) [2/0]  
 | | | marital.status = Married-AF-spouse : <=50K (0/0) [0/0]  
 | | | marital.status = Widowed : <=50K (10.78/1) [7.56/2]  
 | | education = Assoc-acdm : <=50K (22/0) [14.78/1]  
 | | education = Assoc-voc : <=50K (20.45/2) [8.67/1]  
 | | education = 7th-8th : <=50K (5.45/0) [0.78/0]  
 | | education = Doctorate : <=50K (1.11/0) [0.22/0.11]  
 | | education = Prof-school : <=50K (1.22/1.11) [1.11/0]  
 | | education = 5th-6th : <=50K (2.33/0) [0.11/0]  
 | | education = 10th : <=50K (8.78/0) [6.67/0]  
 | | education = 1st-4th : <=50K (0.33/0) [0.11/0]  
 | | education = Preschool : <=50K (0/0) [0.11/0]  
 | | education = 12th : <=50K (6.78/0) [3.44/0]  
 | | occupation = Craft-repair : <=50K (595.5/36.98) [270.46/14.11]  
 | | occupation = Transport-moving : <=50K (221.06/12.36) [106.37/3.04]  
 | | occupation = Farming-fishing : <=50K (123.75/4.2) [64/2.02]  
 | | occupation = Machine-op-inspct : <=50K (336.36/6.55) [149.16/0.06]  
 | | occupation = Tech-support : <=50K (192.51/16.32) [98.67/7.04]  
 | | occupation = Protective-serv : <=50K (95.19/14.16) [50.31/11.02]  
 | | occupation = Armed-Forces : <=50K (4.23/0.01) [0.1/0]  
 | | occupation = Priv-house-serv : <=50K (41.25/0.07) [26/0.01]  
 | capital.gain >= 7599.39 : >50K (230/1) [146/1]  
 relationship = Wife  
 | occupation = Exec-managerial

| | age < 25.21 : <=50K (14.87/2) [5.02/1]  
 | | age >= 25.21  
 | | | education = Bachelors : >50K (68.23/13) [44.48/3.84]  
 | | | education = HS-grad  
 | | | | education.num < 9.06  
 | | | | | capital.gain < 2878  
 | | | | | | workclass = Self-emp-not-inc : <=50K (2.27/0.06) [2.07/1.03]  
 | | | | | | workclass = Private : >50K (26.09/11.44) [15.81/7.49]  
 | | | | | | workclass = State-gov : <=50K (1.13/0.03) [0.04/0.01]  
 | | | | | | workclass = Federal-gov : >50K (0/0) [0/0]  
 | | | | | | workclass = Local-gov : <=50K (2.27/1.06) [2.07/1.03]  
 | | | | | | workclass = Self-emp-inc : >50K (1.13/0.11) [0.04/0.02]  
 | | | | | | workclass = Without-pay : >50K (0/0) [0/0]  
 | | | | | | workclass = Never-worked : <=50K (0/0) [0.2/0]  
 | | | | | capital.gain >= 2878 : >50K (7.41/0) [3.61/1.61]  
 | | | | education.num >= 9.06 : >50K (23/0) [8/0]  
 | | | education = 11th : <=50K (0.2/0) [0/0]  
 | | | education = Masters : >50K (10.82/4.41) [10.41/1.2]  
 | | | education = 9th : <=50K (2.2/1) [0.2/0]  
 | | | education = Some-college : >50K (35.89/11.87) [18.02/7.41]  
 | | | education = Assoc-acdm  
 | | | | fnlwgt < 205216  
 | | | | | age < 36.5 : <=50K (3.61/1) [0.2/0]  
 | | | | | age >= 36.5 : >50K (6/1) [2.2/0]  
 | | | | fnlwgt >= 205216 : >50K (3/0) [2/1]  
 | | | education = Assoc-voc : >50K (10.43/2.82) [2.61/0.2]  
 | | | education = 7th-8th : <=50K (1/0) [0.2/0]  
 | | | education = Doctorate : >50K (2/0) [1/0]  
 | | | education = Prof-school : >50K (3.2/0.2) [2/0]  
 | | | education = 5th-6th : <=50K (0.61/0) [0/0]  
 | | | education = 10th : <=50K (2.41/0) [2.41/0]  
 | | | education = 1st-4th : >50K (0/0) [0/0]  
 | | | education = Preschool : <=50K (0.2/0) [0/0]  
 | | | education = 12th : <=50K (0.2/0.2) [0.41/0]  
 | occupation = Handlers-cleaners  
 | | education = Bachelors : >50K (0.07/0.01) [1.17/0.09]  
 | | education = HS-grad : <=50K (5.26/0.06) [4.11/3.02]  
 | | education = 11th : <=50K (2.04/0) [0.01/0]  
 | | education = Masters : <=50K (0.04/0.02) [0.02/0.01]  
 | | education = 9th : <=50K (0.01/0) [0.01/0]  
 | | education = Some-college : <=50K (1.22/0.05) [1.05/0.03]



| | education = Assoc-acdm : <=50K (0.03/0) [0.02/0.01]  
 | | education = Assoc-voc : <=50K (1.07/0.03) [0.04/0.02]  
 | | education = 7th-8th : <=50K (0/0) [0.01/0]  
 | | education = Doctorate : <=50K (0/0) [0/0]  
 | | education = Prof-school : <=50K (0.01/0) [0/0]  
 | | education = 5th-6th : <=50K (0.03/0) [0.01/0]  
 | | education = 10th : <=50K (1.02/0) [0.02/0]  
 | | education = 1st-4th : <=50K (0/0) [0/0]  
 | | education = Preschool : <=50K (0.01/0) [0/0]  
 | | education = 12th : <=50K (0.01/0.01) [0.02/0]  
 | occupation = Prof-specialty : >50K (240.1/64.15) [143.56/33.9]  
 | occupation = Other-service  
 | | education = Bachelors : <=50K (10.81/4.69) [5.96/1.92]  
 | | education = HS-grad : <=50K (52.12/9.69) [28.27/5.23]  
 | | education = 11th : <=50K (7.46/0) [7.12/2]  
 | | education = Masters : <=50K (0.46/0.23) [0.23/0.12]  
 | | education = 9th : <=50K (2.12/0) [2.12/1]  
 | | education = Some-college : <=50K (20.66/3.58) [17.58/2.35]  
 | | education = Assoc-acdm : <=50K (7.35/3) [0.23/0.12]  
 | | education = Assoc-voc  
 | | | fnlwgt < 115555 : <=50K (2.23/0) [0.12/0]  
 | | | fnlwgt >= 115555 : >50K (3.58/1.23) [2.35/1.12]  
 | | education = 7th-8th : <=50K (5/0) [4.12/0]  
 | | education = Doctorate : <=50K (0/0) [0/0]  
 | | education = Prof-school : <=50K (0.12/0) [0/0]  
 | | education = 5th-6th : <=50K (1.35/0) [0.12/0]  
 | | education = 10th : <=50K (7.23/0) [1.23/0]  
 | | education = 1st-4th : <=50K (3/0) [0/0]  
 | | education = Preschool : <=50K (0.12/0) [0/0]  
 | | education = 12th : <=50K (5.12/0.12) [0.23/0]  
 | occupation = Adm-clerical  
 | | capital.gain < 4225 : <=50K (244.64/106.25) [118.09/51.54]  
 | | capital.gain >= 4225 : >50K (18.18/0) [9.47/1]  
 | occupation = Sales : <=50K (88.69/34.83) [51.9/15.35]  
 | occupation = Craft-repair : <=50K (19.47/9.4) [7.86/2.3]  
 | occupation = Transport-moving  
 | | capital.gain < 4621 : <=50K (14.15/4.26) [1.68/0.22]  
 | | capital.gain >= 4621 : >50K (2.07/-0) [1.03/0]  
 | occupation = Farming-fishing : <=50K (11.9/0.25) [4.52/0.18]  
 | occupation = Machine-op-inspct : <=50K (64.89/15.34) [26.85/3.99]  
 | occupation = Tech-support

| | education = Bachelors : >50K (6.2/0.03) [4.48/0.25]

| | education = HS-grad

| | | age < 32.5 : <=50K (2.34/0.03) [0.14/0.03]

| | | age >= 32.5 : >50K (3.42/0.28) [4.17/3.14]

| | education = 11th : <=50K (0.11/0) [0.03/0]

| | education = Masters : <=50K (0.11/0.06) [0.06/0.03]

| | education = 9th : <=50K (0.03/0) [1.03/0]

| | education = Some-college : <=50K (8.65/3.14) [5.14/3.08]

| | education = Assoc-acdm : <=50K (3.08/0) [0.06/0.03]

| | education = Assoc-voc : >50K (5.2/1.11) [6.11/1.06]

| | education = 7th-8th : <=50K (0/0) [0.03/0]

| | education = Doctorate : <=50K (1/0) [0/0]

| | education = Prof-school : <=50K (0.03/0) [0/0]

| | education = 5th-6th : <=50K (0.08/0) [0.03/0]

| | education = 10th : <=50K (0.06/0) [0.06/0]

| | education = 1st-4th : >50K (0/0) [0/0]

| | education = Preschool : <=50K (0.03/0) [0/0]

| | education = 12th : >50K (1.03/0) [0.06/0.06]

| occupation = Protective-serv : <=50K (5.41/1.11) [2.24/1.08]

| occupation = Armed-Forces : >50K (0/0) [0/0]

| occupation = Priv-house-serv : <=50K (5.41/0.11) [4.24/0.08]

relationship = Own-child

| capital.gain < 4718.5 : <=50K (3391/32) [1646/14]

| capital.gain >= 4718.5 : >50K (24/9) [7/1]

relationship = Unmarried

| capital.gain < 7139.5 : <=50K (2213/108) [1177/58]

| capital.gain >= 7139.5 : >50K (42/3) [14/1]

relationship = Other-relative

| native.country = United-States

| | capital.gain < 5699.5

| | | capital.loss < 2113

| | | | marital.status = Married-civ-spouse

| | | | | education = Bachelors : <=50K (10/1) [3/0]

| | | | | education = HS-grad : <=50K (12/0) [16/2]

| | | | | education = 11th : <=50K (2/0) [3/0]

| | | | | education = Masters : <=50K (1/0) [0/0]

| | | | | education = 9th : <=50K (0/0) [0/0]

| | | | | education = Some-college

| | | | | occupation = Exec-managerial : >50K (2/1) [1/0]

| | | | | occupation = Handlers-cleaners : <=50K (0/0) [0/0]

| | | | | occupation = Prof-specialty : <=50K (1/0) [0/0]

| | | | | | occupation = Other-service : <=50K (1/0) [0/0]  
 | | | | | | occupation = Adm-clerical : >50K (3/1) [1/0]  
 | | | | | | occupation = Sales : <=50K (2/0) [0/0]  
 | | | | | | occupation = Craft-repair : <=50K (0/0) [0/0]  
 | | | | | | occupation = Transport-moving : <=50K (0/0) [0/0]  
 | | | | | | occupation = Farming-fishing : <=50K (0/0) [0/0]  
 | | | | | | occupation = Machine-op-inspct : <=50K (0/0) [0/0]  
 | | | | | | occupation = Tech-support : <=50K (0/0) [0/0]  
 | | | | | | occupation = Protective-serv : <=50K (1/0) [0/0]  
 | | | | | | occupation = Armed-Forces : <=50K (0/0) [0/0]  
 | | | | | | occupation = Priv-house-serv : <=50K (0/0) [0/0]  
 | | | | | | education = Assoc-acdm : <=50K (0/0) [0/0]  
 | | | | | | education = Assoc-voc : <=50K (2/0) [3/1]  
 | | | | | | education = 7th-8th : <=50K (1.73/0) [0/0]  
 | | | | | | education = Doctorate : <=50K (0/0) [0/0]  
 | | | | | | education = Prof-school : <=50K (0/0) [0/0]  
 | | | | | | education = 5th-6th : <=50K (1/0) [1/0]  
 | | | | | | education = 10th : <=50K (0/0) [1/0]  
 | | | | | | education = 1st-4th : <=50K (2/0) [0/0]  
 | | | | | | education = Preschool : <=50K (0/0) [0/0]  
 | | | | | | education = 12th : <=50K (1/0) [0/0]  
 | | | | marital.status = Divorced : <=50K (54.73/2) [39/1]  
 | | | | marital.status = Married-spouse-absent : <=50K (7.73/0) [5.19/0]  
 | | | | marital.status = Never-married : <=50K (274.2/1) [174.1/0]  
 | | | | marital.status = Separated : <=50K (22/1) [18/0]  
 | | | | marital.status = Married-AF-spouse : <=50K (1/0) [0/0]  
 | | | | marital.status = Widowed : <=50K (26.73/1) [8/0]  
 | | | capital.loss >= 2113 : >50K (5/1) [0/0]  
 | | capital.gain >= 5699.5 : >50K (6.73/1) [3/0]  
 | native.country = Cuba : <=50K (3.1/0.01) [2.05/0]  
 | native.country = Jamaica : <=50K (3.1/0.01) [6.05/0]  
 | native.country = India : <=50K (1.03/0) [5.02/1]  
 | native.country = Mexico : <=50K (62.98/0.1) [31.04/0]  
 | native.country = South : <=50K (7.23/0.01) [2.12/0]  
 | native.country = Puerto-Rico : <=50K (5.16/0.01) [3.09/0]  
 | native.country = Honduras : <=50K (1.03/0) [0.02/0]  
 | native.country = England : <=50K (1.03/0) [0.02/0]  
 | native.country = Canada : >50K (1.03/0.03) [0.02/0.02]  
 | native.country = Germany : <=50K (1.03/0) [1.02/0]  
 | native.country = Iran : >50K (1.03/0.03) [0.02/0.02]  
 | native.country = Philippines : <=50K (12.39/1.02) [8.2/1]

| native.country = Italy : <=50K (1.03/0) [1.02/0]  
 | native.country = Poland : <=50K (3.1/0.01) [1.05/0]  
 | native.country = Columbia : <=50K (1.03/0) [2.02/0]  
 | native.country = Cambodia : <=50K (2.06/1) [1.03/0]  
 | native.country = Thailand : <=50K (1.03/0) [0.02/0]  
 | native.country = Ecuador : <=50K (4.13/1.01) [3.07/0]  
 | native.country = Laos : <=50K (0/0) [0/0]  
 | native.country = Taiwan : <=50K (1.03/0) [0.02/0]  
 | native.country = Haiti : <=50K (6.19/0.01) [4.1/0]  
 | native.country = Portugal : <=50K (0/0) [0/0]  
 | native.country = Dominican-Republic : <=50K (4.13/0.01) [2.07/0]  
 | native.country = El-Salvador : <=50K (9.29/0.02) [9.15/0]  
 | native.country = France : <=50K (0/0) [0/0]  
 | native.country = Guatemala : <=50K (7.23/1.01) [6.12/0]  
 | native.country = China : <=50K (2.06/1) [3.03/0]  
 | native.country = Japan : <=50K (1.03/0) [0.02/0]  
 | native.country = Yugoslavia : <=50K (0/0) [0/0]  
 | native.country = Peru : <=50K (2.06/0) [0.03/0]  
 | native.country = Outlying-US(Guam-USVI-etc) : <=50K (0/0) [0/0]  
 | native.country = Scotland : <=50K (0/0) [0/0]  
 | native.country = Trinidad&Tobago : <=50K (1.03/0) [0.02/0]  
 | native.country = Greece : <=50K (0/0) [1/0]  
 | native.country = Nicaragua : <=50K (5.16/0.01) [1.09/0]  
 | native.country = Vietnam : <=50K (10.32/0.02) [2.17/0]  
 | native.country = Hong : <=50K (0/0) [3/0]  
 | native.country = Ireland : <=50K (1.03/0) [0.02/0]  
 | native.country = Hungary : <=50K (1.03/0) [0.02/0]  
 | native.country = Holand-Netherlands : <=50K (0/0) [1/0]

### 5.3 Modello di Silhouette con il Naive Bayes

Attribute	opel	saab	bus	van
	(0.25)	(0.26)	(0.26)	(0.24)

---

#### COMPACTNESS

mean	94.9974	97.2453	91.578	90.5056
std. dev.	8.1774	9.0909	8.6065	4.0111
weight sum	212	217	218	199
precision	1.0698	1.0698	1.0698	1.0698

#### CIRCULARITY

mean	46.5802	45.5346	45.0688	42.0704
std. dev.	7.2182	6.8034	5.0192	4.0728
weight sum	212	217	218	199
precision	1	1	1	1

#### DISTANCE\_CIRCULARITY

mean	89.0743	88.6541	76.7197	73.3189
std. dev.	15.5555	17.026	12.0292	10.8137
weight sum	212	217	218	199
precision	1.1613	1.1613	1.1613	1.1613

#### RADIUS\_RATIO

mean	180.2859	180.853	166.0278	147.0802
std. dev.	31.2924	30.7223	30.4572	29.8102
weight sum	212	217	218	199
precision	1.7218	1.7218	1.7218	1.7218

#### PR.AXIS\_ASPECT\_RATIO

mean	61.0721	61.2258	63.4264	61.2891
std. dev.	5.0572	4.3862	8.8268	11.4274
weight sum	212	217	218	199
precision	2.5278	2.5278	2.5278	2.5278

#### MAX.LENGTH\_ASPECT\_RATIO

mean	8.925	8.8903	7.1842	9.6412
std. dev.	2.0855	2.2589	4.8718	7.1185
weight sum	212	217	218	199
precision	2.65	2.65	2.65	2.65

#### SCATTER\_RATIO

mean	182.1733	179.695	169.979	141.5383
std. dev.	32.7443	31.4383	33.2899	13.9536
weight sum	212	217	218	199
precision	1.1769	1.1769	1.1769	1.1769

#### ELONGATEDNESS

mean	37.9134	38.3681	40.2651	48.1289
std. dev.	7.8238	7.5772	6.6772	4.493
weight sum	212	217	218	199
precision	1.0294	1.0294	1.0294	1.0294

#### PR.AXIS\_RECTANGULARITY

mean	21.5896	21.447	20.5734	18.5779
std. dev.	2.547	2.4473	2.7221	1.0286
weight sum	212	217	218	199
precision	1	1	1	1

#### MAX.LENGTH\_RECTANGULARITY

mean	151.2772	148.6898	146.7876	145.2006
std. dev.	18.0956	16.0797	10.4646	11.0053
weight sum	212	217	218	199
precision	1.0769	1.0769	1.0769	1.0769

#### SCALED\_VARIANCE\_MAJOR

mean	198.5741	197.1011	192.9029	164.1234
std. dev.	28.6335	27.7863	33.8236	19.6532
weight sum	212	217	218	199
precision	1.4961	1.4961	1.4961	1.4961

#### SCALED\_VARIANCE\_MINOR

mean	508.4762	493.7891	448.8265	298.2018
std. dev.	172.0663	162.7785	192.7775	55.8744
weight sum	212	217	218	199
precision	1.9716	1.9716	1.9716	1.9716

#### SCALED\_RADIUS\_OF\_GYRATION

mean	179.7782	179.3974	180.9989	157.2501
std. dev.	34.684	33.6916	31.1784	22.7642
weight sum	212	217	218	199
precision	1.1197	1.1197	1.1197	1.1197

#### SKEWNESS\_ABOUT\_MAJOR

mean	70.2358	69.8065	77.055	72.8342
std. dev.	5.1799	5.3088	7.7345	8.8634
weight sum	212	217	218	199
precision	2	2	2	2

#### SKEWNESS\_ABOUT\_MINOR

mean	6.6038	7.659	4.844	6.4171
std. dev.	5.181	5.8035	3.2131	4.6541
weight sum	212	217	218	199
precision	1	1	1	1

#### KURTOSIS\_ABOUT\_MAJOR

mean	15.0462	15.3419	10.3769	9.8791
std. dev.	10.0009	9.9013	6.8491	6.2513
weight sum	212	217	218	199
precision	1.025	1.025	1.025	1.025

#### KURTOSIS\_ABOUT\_MINOR

mean	189.4275	189.8729	187.8061	189.0348
std. dev.	5.4966	4.8815	7.1897	6.259
weight sum	212	217	218	199
precision	1.0345	1.0345	1.0345	1.0345

#### HOLLOWS\_RATIO

mean	197.1132	198.0415	191.3257	196.1457
std. dev.	5.8323	6.5951	7.9001	7.3268
weight sum	212	217	218	199
precision	1	1	1	1

### 5.4 Modello di Decision Tree su Silhouette

```
ELONGATEDNESS < 41.5
|  MAX.LENGTH_ASPECT_RATIO < 7.5
| |  COMPACTNESS < 95.5
| | |  PR.AXIS_ASPECT_RATIO < 67.5
| | | |  SCATTER_RATIO < 161.5 : saab (4/0) [1/0]
```

```

| | | SCATTER_RATIO >= 161.5 : opel (13/6) [3/1]
| | | PR.AXIS_ASPECT_RATIO >= 67.5 : bus (13/0) [3/0]
| | COMPACTNESS >= 95.5 : bus (52/0) [18/1]
| MAX.LENGTH_ASPECT_RATIO >= 7.5
| | SCALED_VARIANCE_MINOR < 721
| | | HOLLOWS_RATIO < 195.5 : opel (27/4) [12/3]
| | | HOLLOWS_RATIO >= 195.5
| | | COMPACTNESS < 109.5
| | | | DISTANCE_CIRCULARITY < 99 : opel (56/26) [26/12]
| | | | DISTANCE_CIRCULARITY >= 99
| | | | KURTOSIS_ABOUT_MINOR < 187.5 : saab (14/0) [5/2]
| | | | KURTOSIS_ABOUT_MINOR >= 187.5
| | | | | HOLLOWS_RATIO < 204.5
| | | | | SKEWNESS_ABOUT_MAJOR < 66.5 : opel (13/2) [2/0]
| | | | | SKEWNESS_ABOUT_MAJOR >= 66.5 : saab (36/11) [29/10]
| | | | | HOLLOWS_RATIO >= 204.5 : saab (8/0) [7/4]
| | | COMPACTNESS >= 109.5 : saab (13/0) [3/0]
| | SCALED_VARIANCE_MINOR >= 721 : opel (16/1) [8/1]
| ELONGATEDNESS >= 41.5
| | MAX.LENGTH_ASPECT_RATIO < 8.5
| | | SCALED_VARIANCE_MINOR < 298.5
| | | | MAX.LENGTH_RECTANGULARITY < 131.5
| | | | | PR.AXIS_RECTANGULARITY < 17.5 : van (20/8) [10/4]
| | | | | PR.AXIS_RECTANGULARITY >= 17.5
| | | | | SCALED_RADIUS_OF_GYRATION < 127.5
| | | | | COMPACTNESS < 88.5
| | | | | | DISTANCE_CIRCULARITY < 68.5 : van (2/0) [1/0]
| | | | | | DISTANCE_CIRCULARITY >= 68.5 : opel (2/0) [1/0]
| | | | | COMPACTNESS >= 88.5 : saab (4/1) [2/2]
| | | | | SCALED_RADIUS_OF_GYRATION >= 127.5
| | | | | SCALED_RADIUS_OF_GYRATION < 136 : opel (4/0) [4/2]
| | | | | SCALED_RADIUS_OF_GYRATION >= 136
| | | | | CIRCULARITY < 37.5 : saab (12/2) [2/0]
| | | | | CIRCULARITY >= 37.5 : opel (8/2) [1/0]
| | | MAX.LENGTH_RECTANGULARITY >= 131.5 : van (54/8) [24/8]
| | SCALED_VARIANCE_MINOR >= 298.5
| | | DISTANCE_CIRCULARITY < 76.5
| | | | SKEWNESS_ABOUT_MINOR < 10.5 : bus (85/11) [58/9]
| | | | SKEWNESS_ABOUT_MINOR >= 10.5
| | | | | PR.AXIS_ASPECT_RATIO < 62.5 : opel (12/4) [5/2]
| | | | | PR.AXIS_ASPECT_RATIO >= 62.5 : bus (2/0) [1/0]
| | | DISTANCE_CIRCULARITY >= 76.5 : opel (21/9) [13/8]
| MAX.LENGTH_ASPECT_RATIO >= 8.5 : van (73/4) [43/6]

```