



Laurea Triennale in informatica - Università di Salerno
Corso di *Machine Learning* - Prof.ssa L. Caruccio, Prof. G. Polese



Machine Learning

PROJECT

Allegati Progetto Machine Learning

Versione	1.0
Data	16/02/2024
Destinatario	Dipartimento di Informatica dell'Università degli studi di Salerno
Presentato da	Cristian Porzio, Raffaele Monti



Revision History

Data	Versione	Descrizione	Autori
16/02/2024	1.0	Stesura e Revisione completa	Raffaele Monti, Cristian Porzio

Ruoli e Contatti

Ruolo	Nome	Contatti
Docente	Giuseppe Polese	gpolese@unisa.it
Docente	Loredana Caruccio	lcaruccio@unisa.it
Studente	Raffaele Monti	r.monti2@studenti.unisa.it
Studente	Cristian Porzio	c.porzio3@studenti.unisa.it



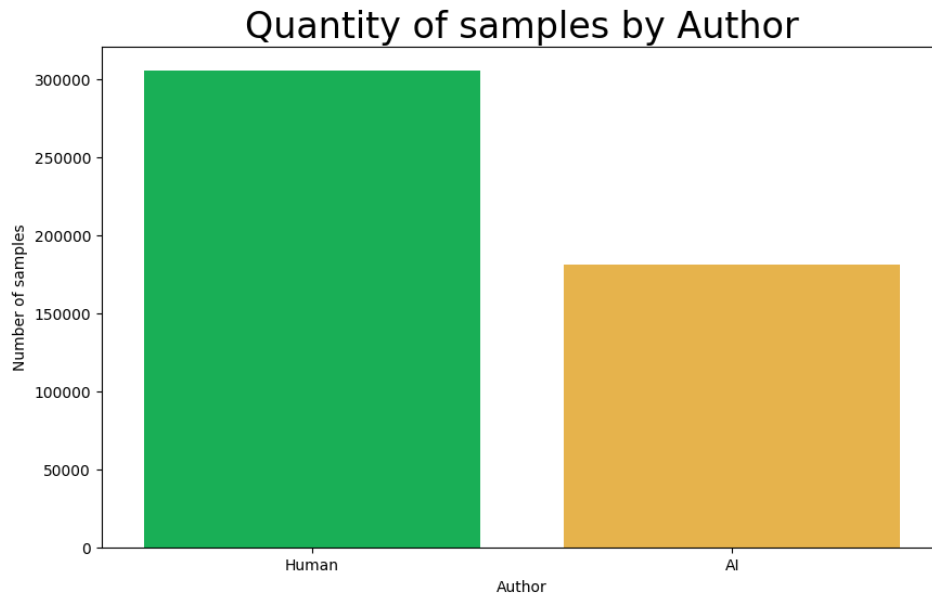
Sommario

1. Dataset 1 - Stato iniziale	4
2. Dataset 1 non bilanciato – Esecuzione.....	7
3. Dataset 1 bilanciato – Esecuzione	10
4. Dataset 2 - Stato iniziale	13
5. Dataset 2 non bilanciato – Esecuzione.....	16
6. Dataset 2 bilanciato – Esecuzione	19
7. DTC varie profondità – Dataset 1 non bilanciato	22
8. DTC varie profondità – Dataset 1 bilanciato.....	28
9. DTC varie profondità – Dataset 2 non bilanciato	34
10. DTC varie profondità – Dataset 2 bilanciato	40

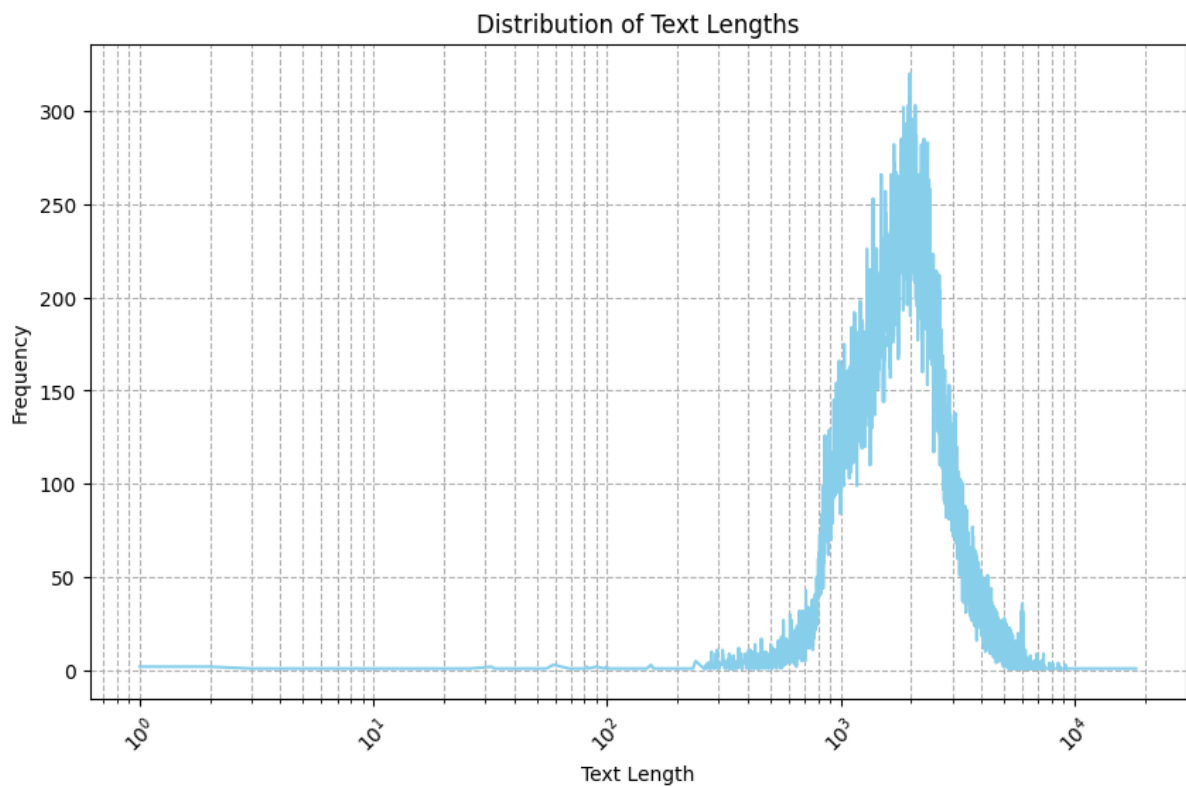


1. Dataset 1 - Stato iniziale

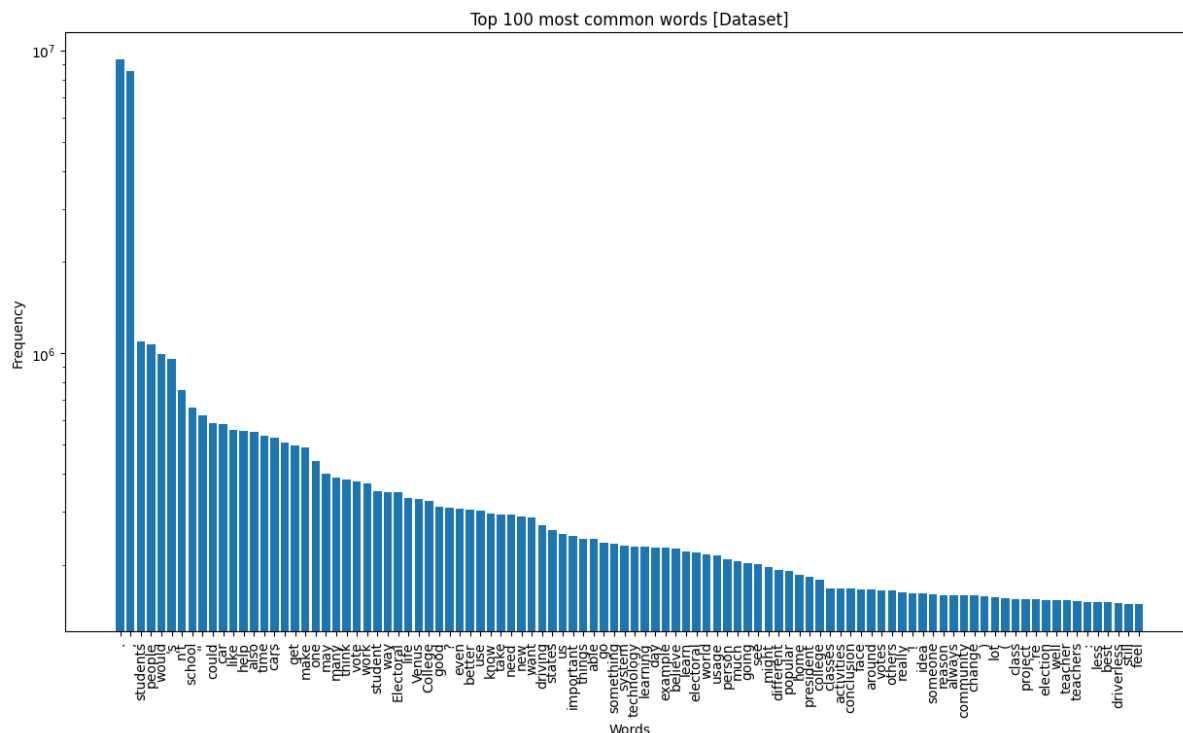
1.1 Quantità di testi per autore



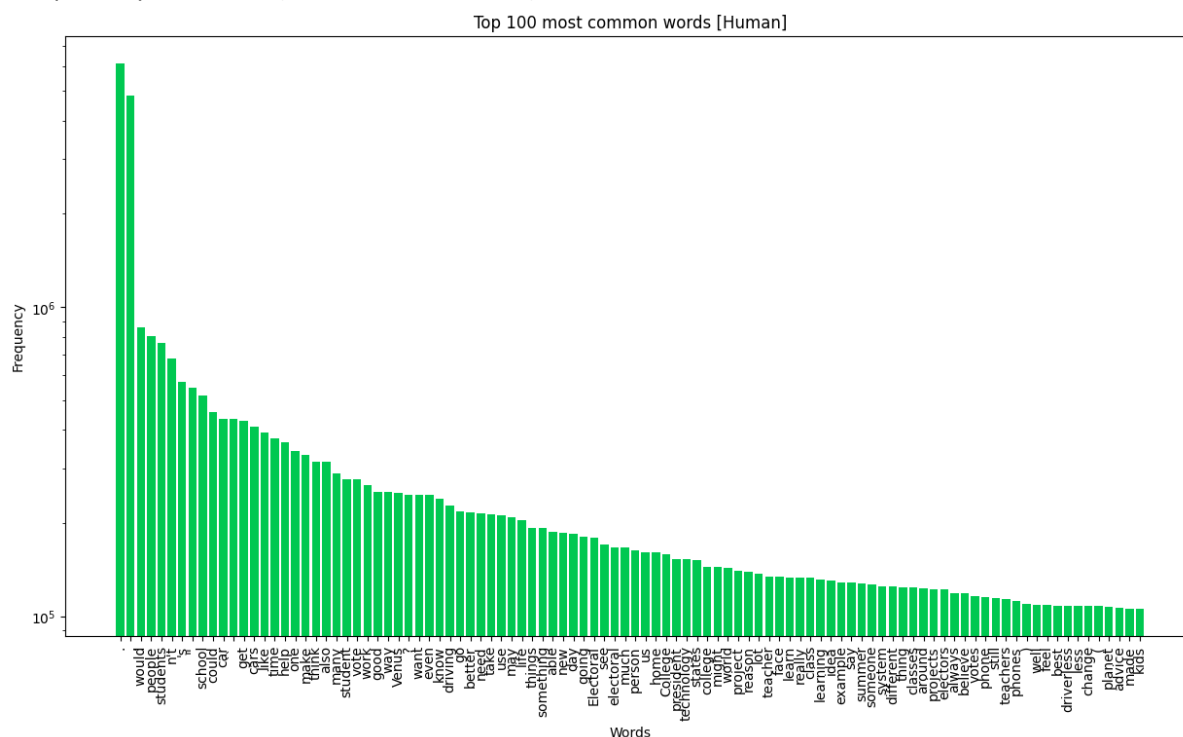
1.2 Distribuzione della lunghezza dei testi



1.3 100 parole più comuni (Intero Dataset 1)

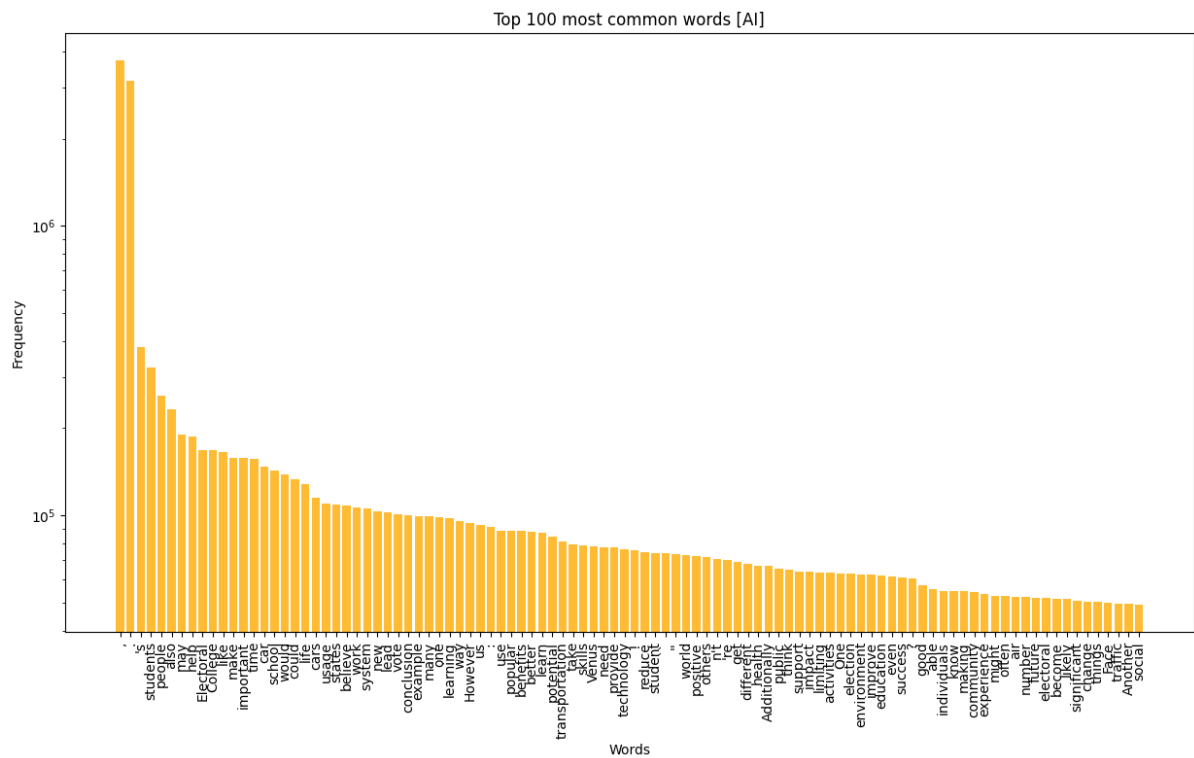


1.4 100 parole più comuni (solo autore 'Human')





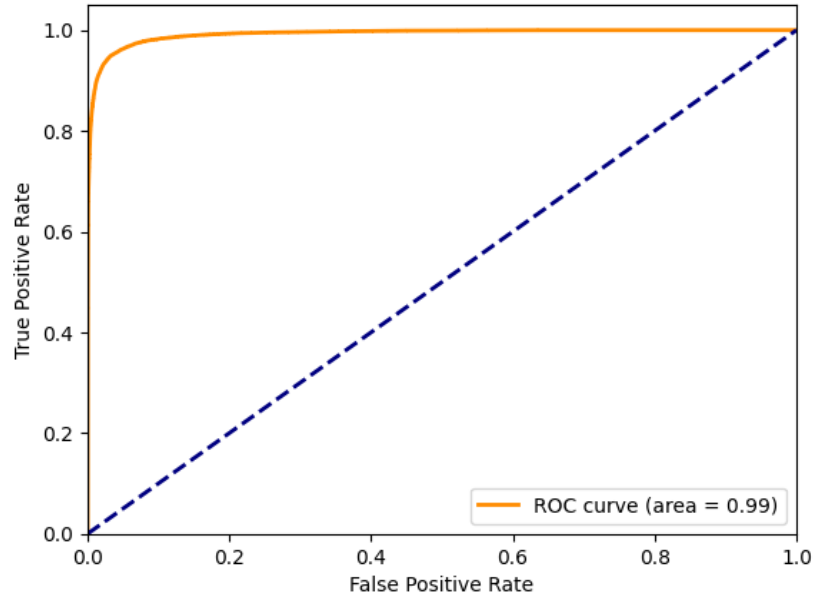
1.5 100 parole più comuni (solo autore 'AI')



2. Dataset 1 non bilanciato – Esecuzione

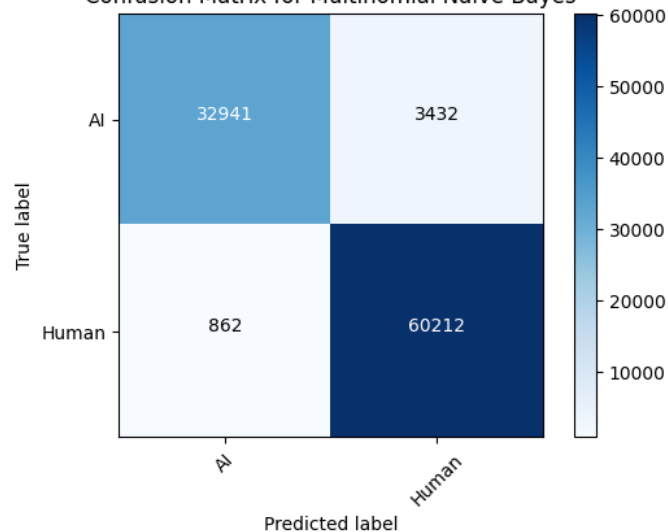
2.1 ROC modello M-NB

Receiver Operating Characteristic (ROC) Curve for Multinomial Naive Bayes



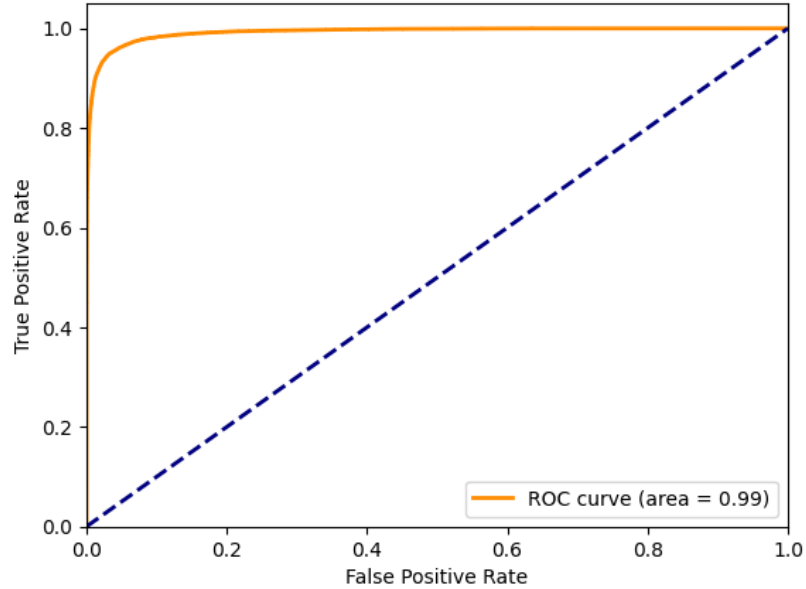
2.2 Confusion Matrix modello M-NB

Confusion Matrix for Multinomial Naive Bayes



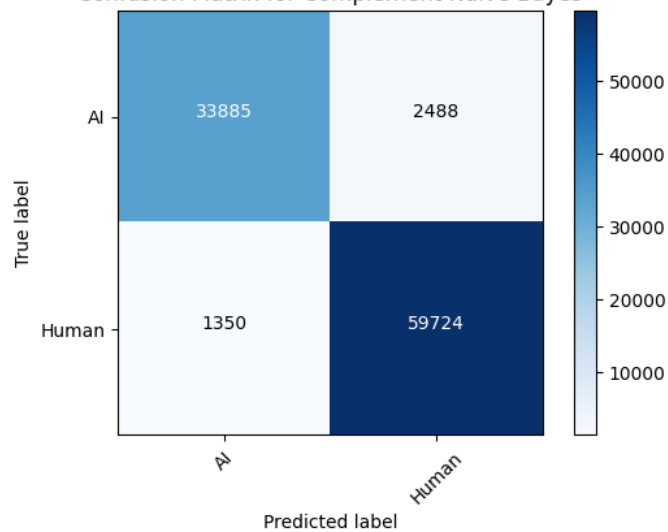
2.3 ROC modello C-NB

Receiver Operating Characteristic (ROC) Curve for Complement Naive Bayes



2.4 Confusion Matrix modello C-NB

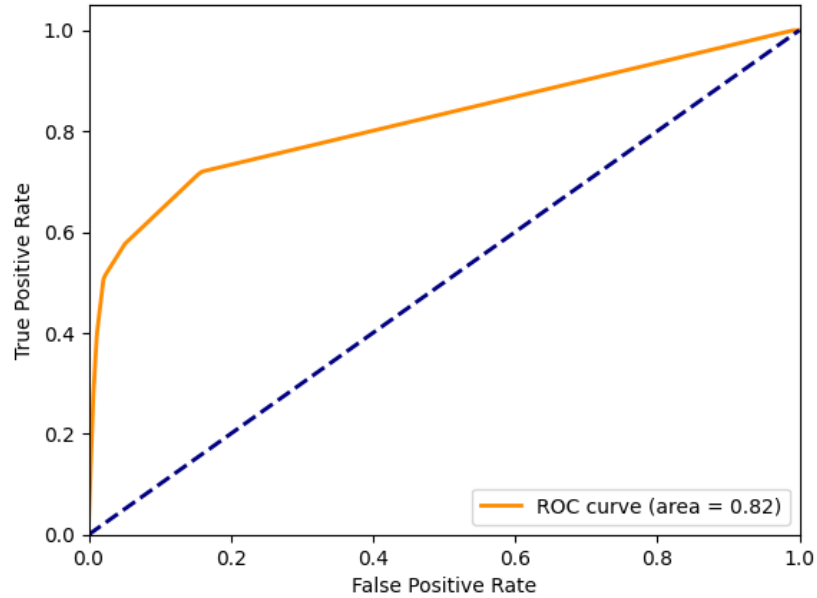
Confusion Matrix for Complement Naive Bayes





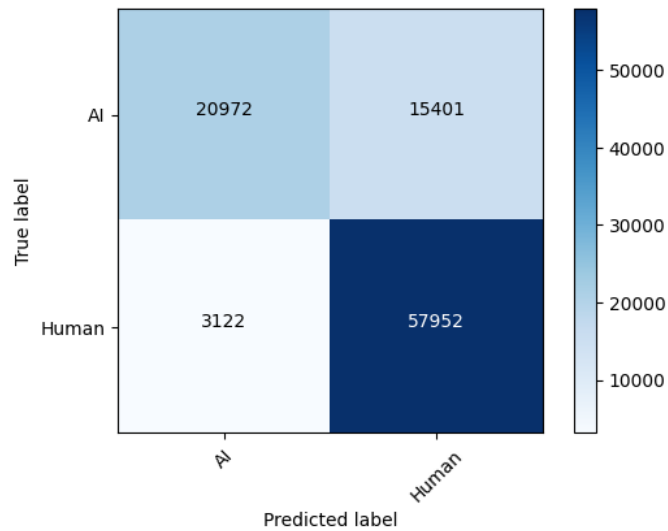
2.5 ROC modello DTC

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier



2.6 Confusion Matrix modello DTC

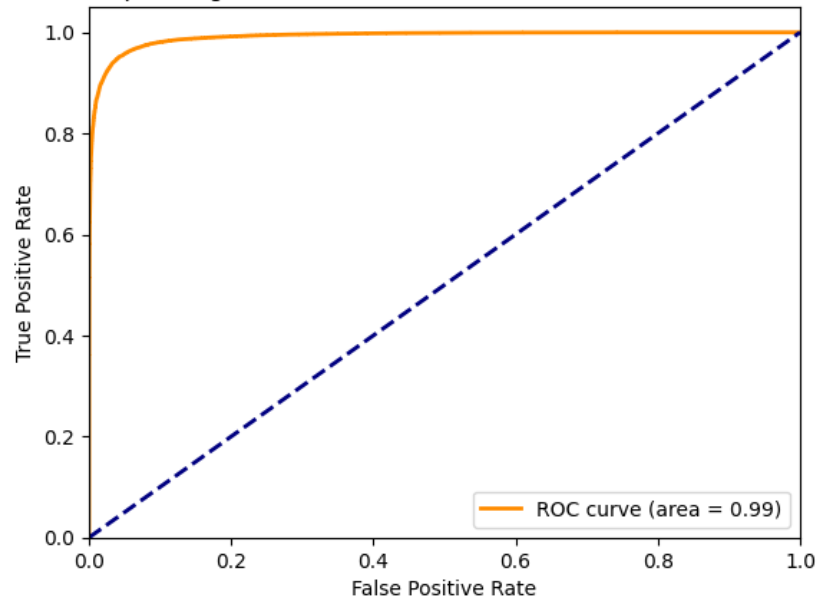
Confusion Matrix for Decision Tree Classifier



3. Dataset 1 bilanciato – Esecuzione

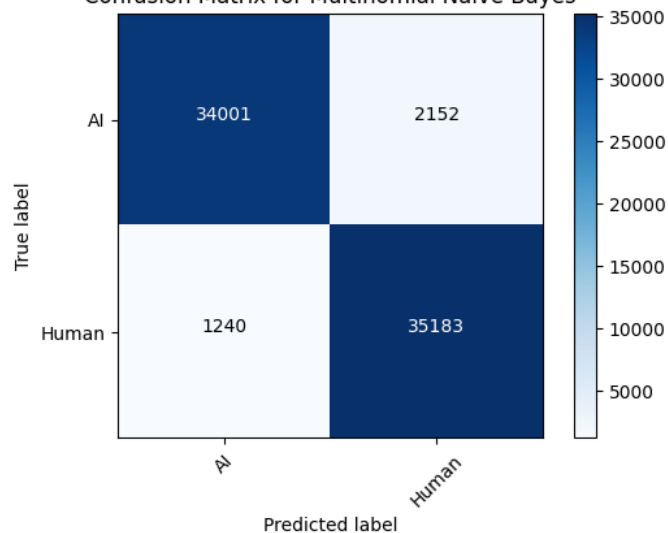
3.1 ROC modello M-NB

Receiver Operating Characteristic (ROC) Curve for Multinomial Naive Bayes



3.2 Confusion Matrix modello M-NB

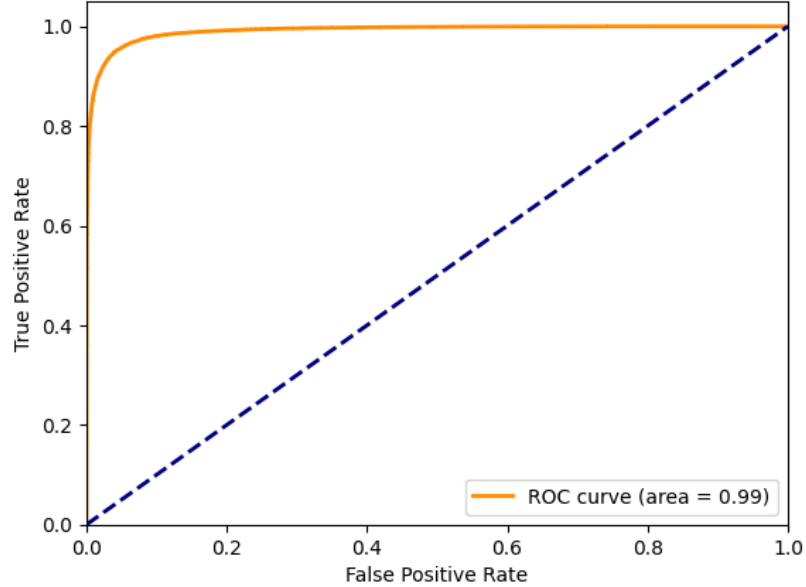
Confusion Matrix for Multinomial Naive Bayes



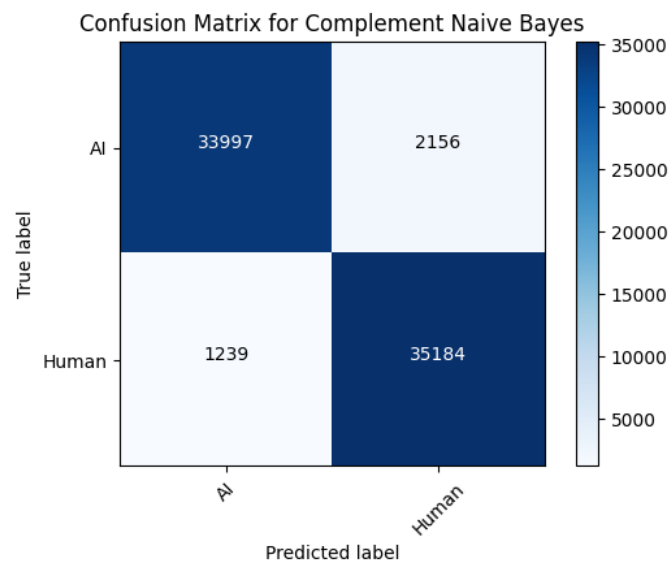


3.3 ROC modello C-NB

Receiver Operating Characteristic (ROC) Curve for Complement Naive Bayes



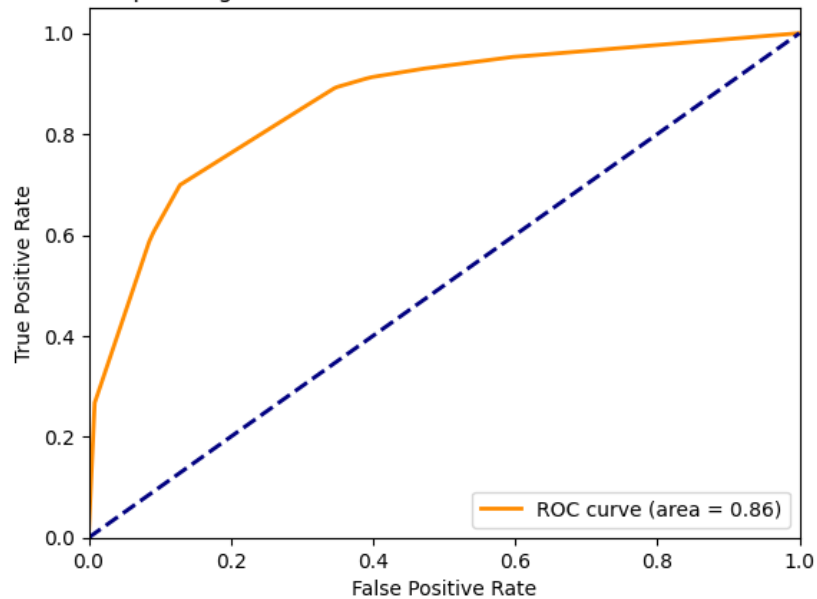
3.4 Confusion Matrix modello C-NB



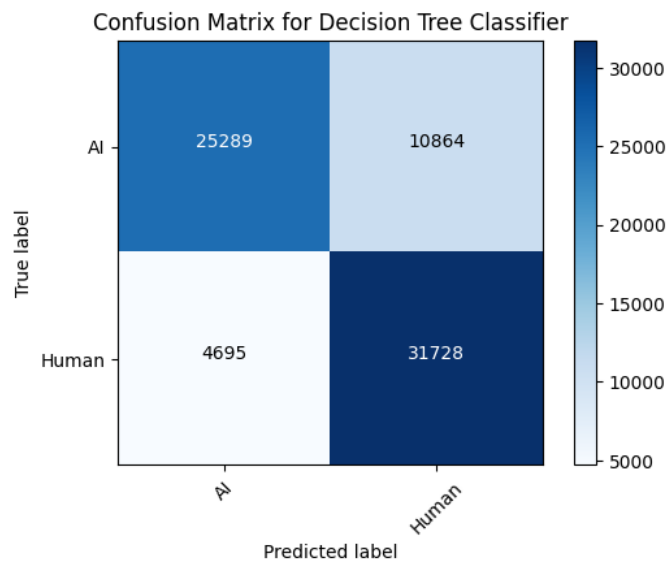


3.5 ROC modello DTC

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier



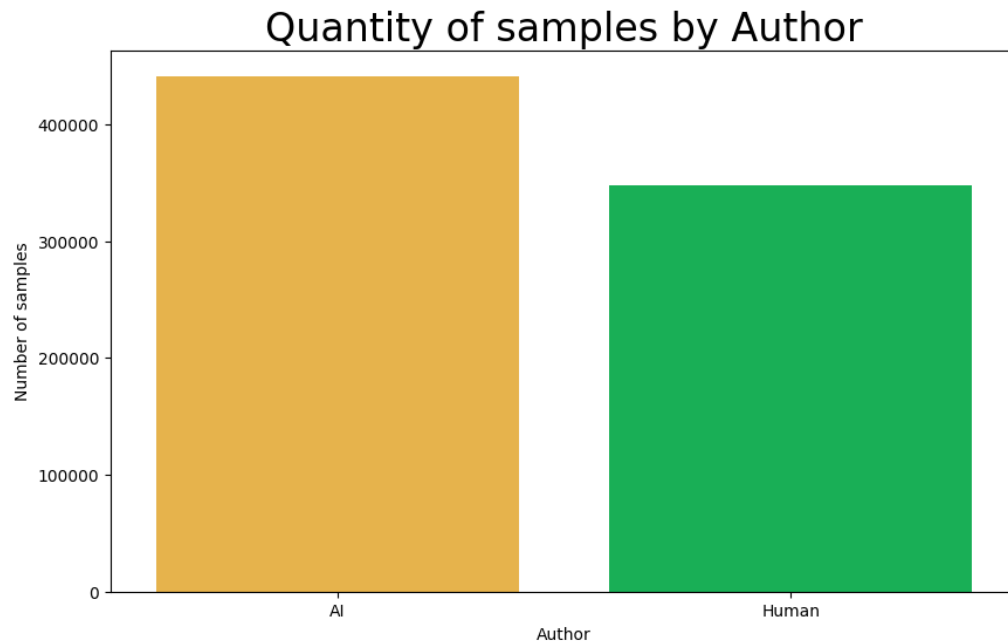
3.6 Confusion Matrix modello DTC



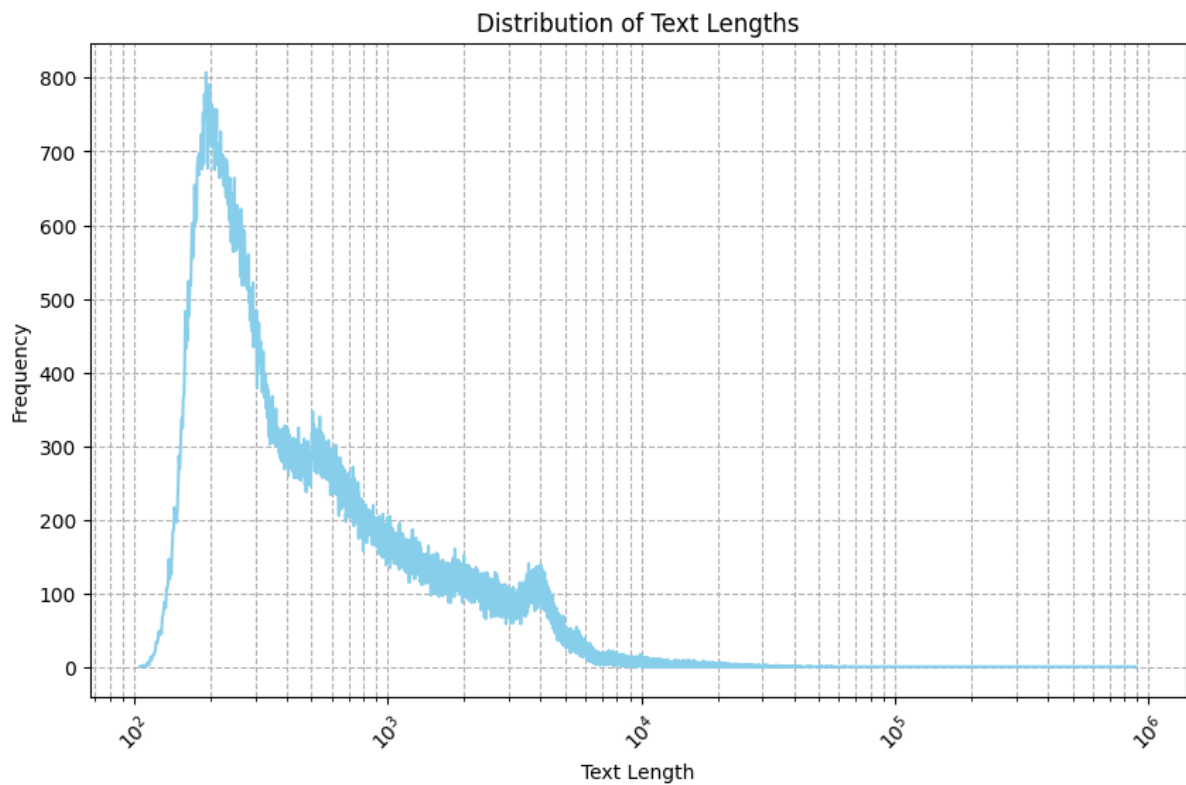


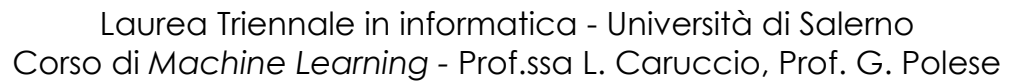
4. Dataset 2 - Stato iniziale

4.1 Quantità di testi per autore

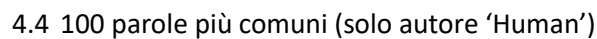


4.2 Distribuzione della lunghezza dei testi



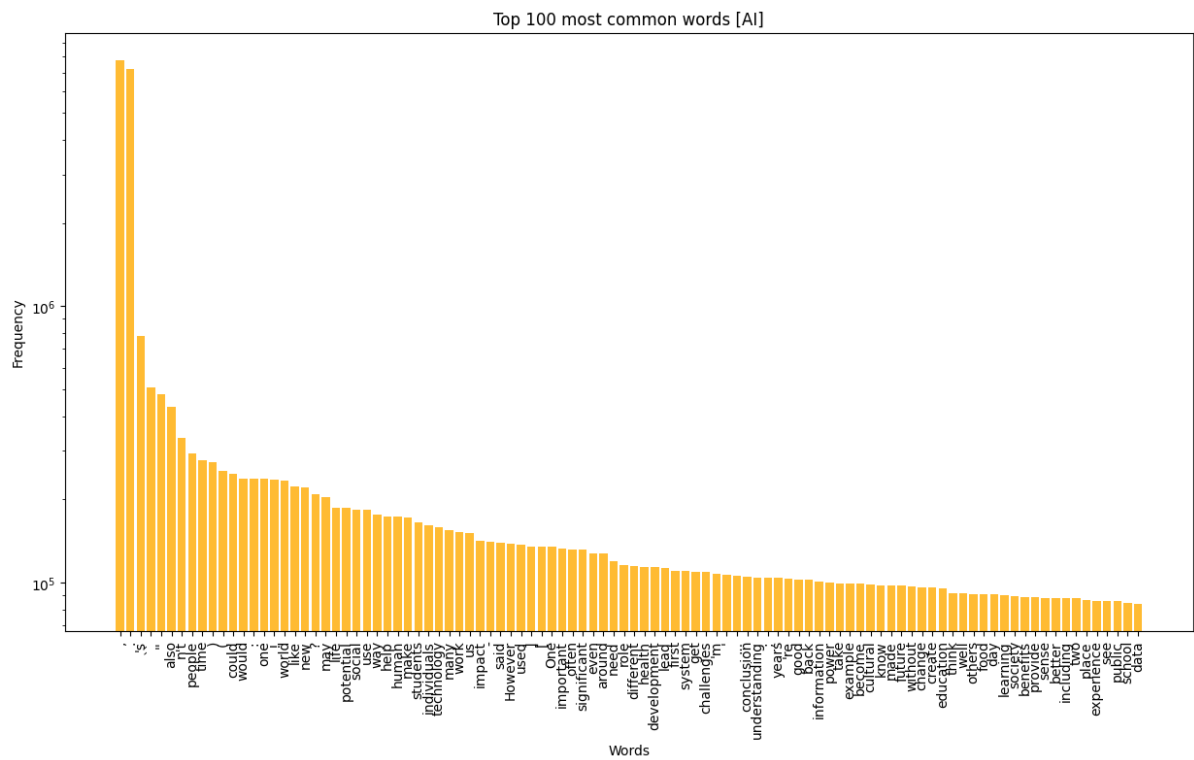


Top 100 most common words [Dataset]





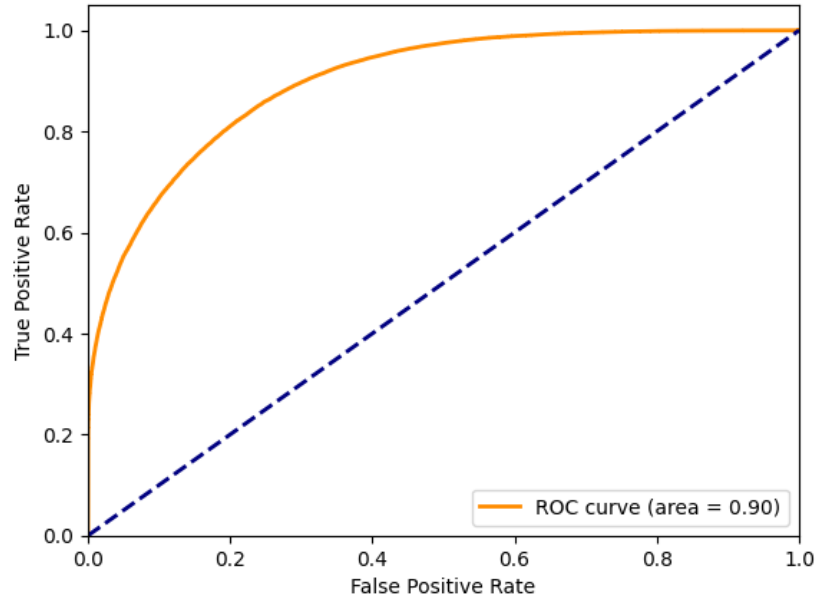
4.5 100 parole più comuni (solo autore 'AI')



5. Dataset 2 non bilanciato – Esecuzione

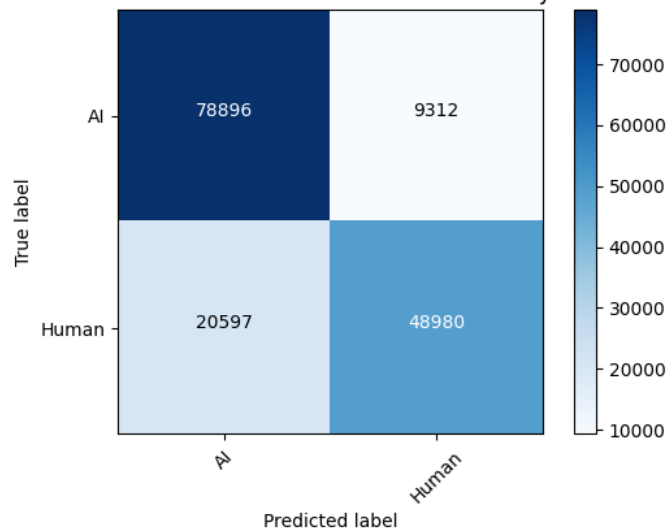
5.1 ROC modello M-NB

Receiver Operating Characteristic (ROC) Curve for Multinomial Naive Bayes



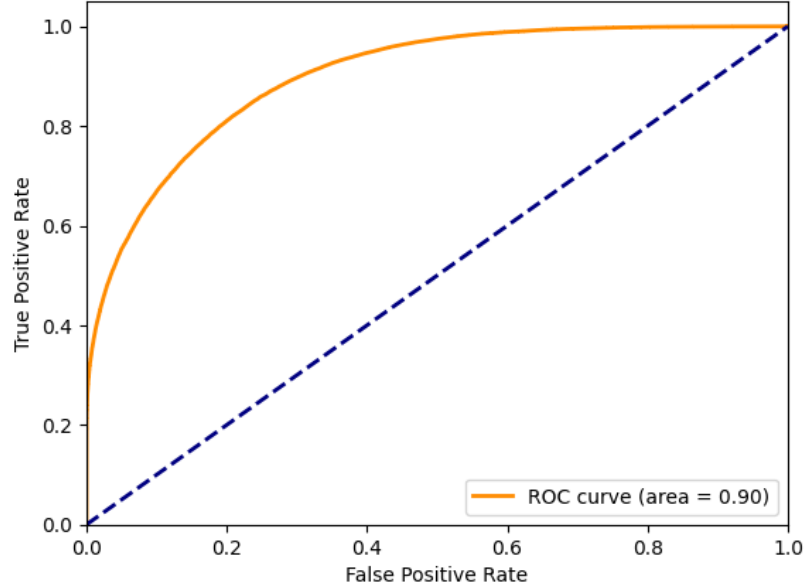
5.2 Confusion Matrix modello M-NB

Confusion Matrix for Multinomial Naive Bayes



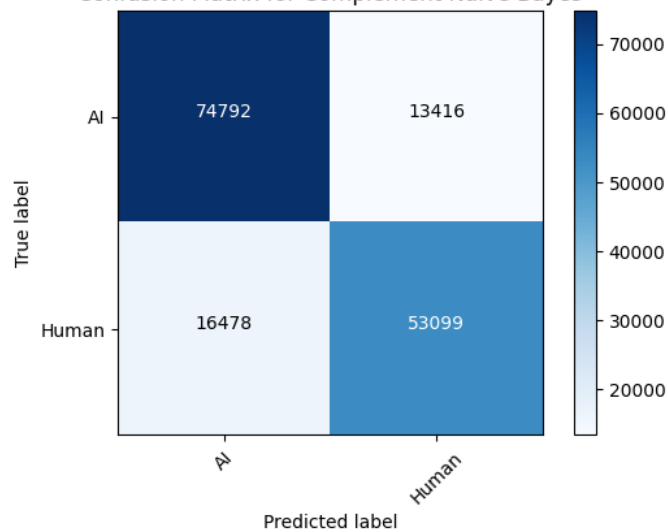
5.3 ROC modello C-NB

Receiver Operating Characteristic (ROC) Curve for Complement Naive Bayes



5.4 Confusion Matrix modello C-NB

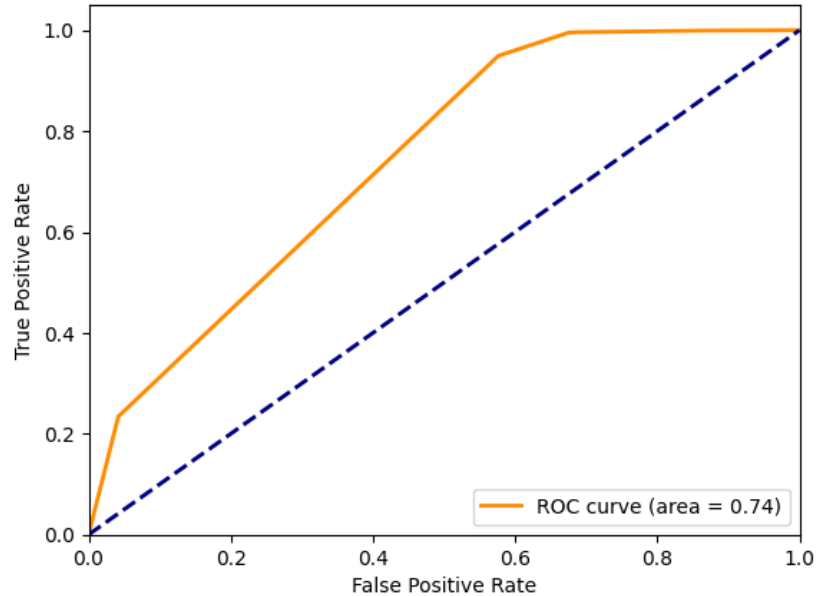
Confusion Matrix for Complement Naive Bayes





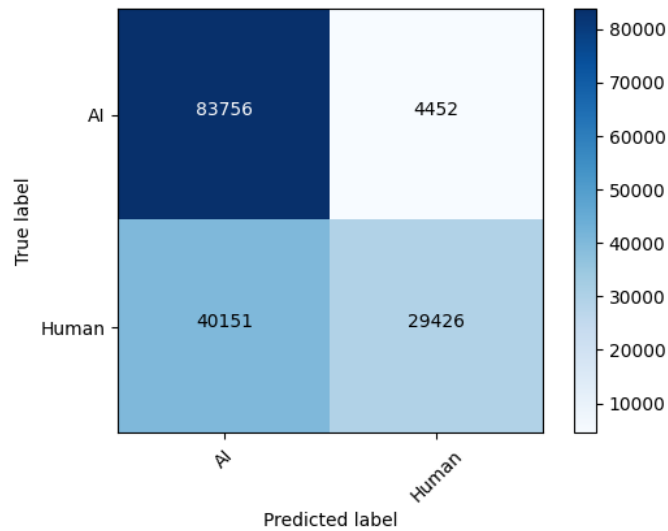
5.5 ROC modello DTC

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier



5.6 Confusion Matrix modello DTC

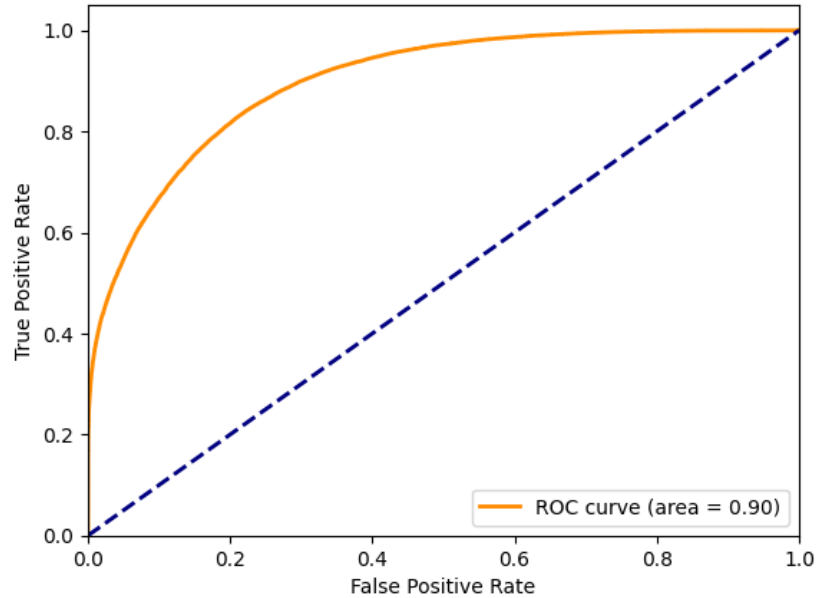
Confusion Matrix for Decision Tree Classifier



6. Dataset 2 bilanciato – Esecuzione

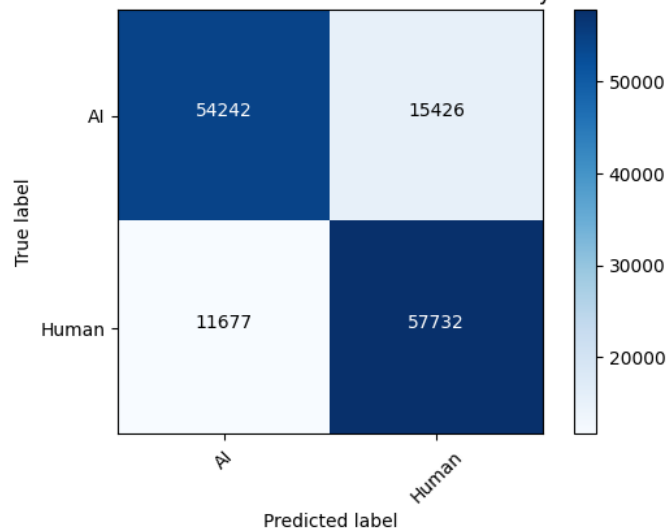
6.1 ROC modello M-NB

Receiver Operating Characteristic (ROC) Curve for Multinomial Naive Bayes



6.2 Confusion Matrix modello M-NB

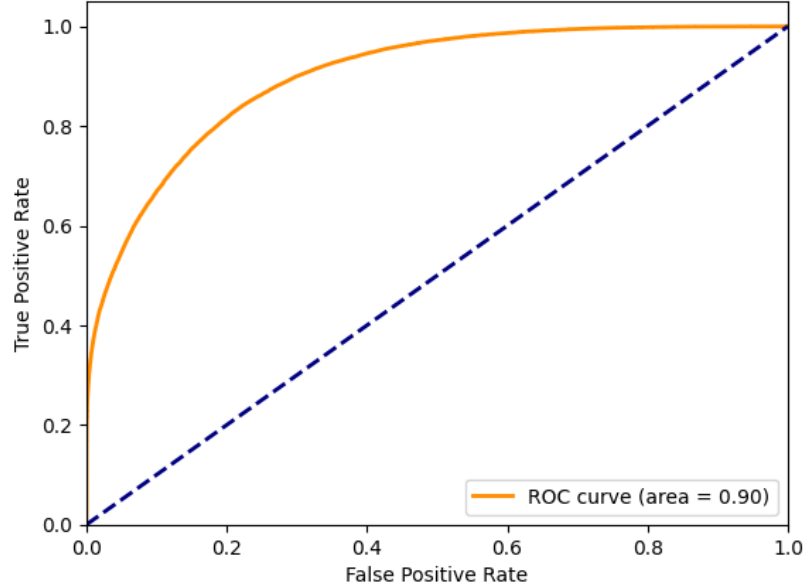
Confusion Matrix for Multinomial Naive Bayes





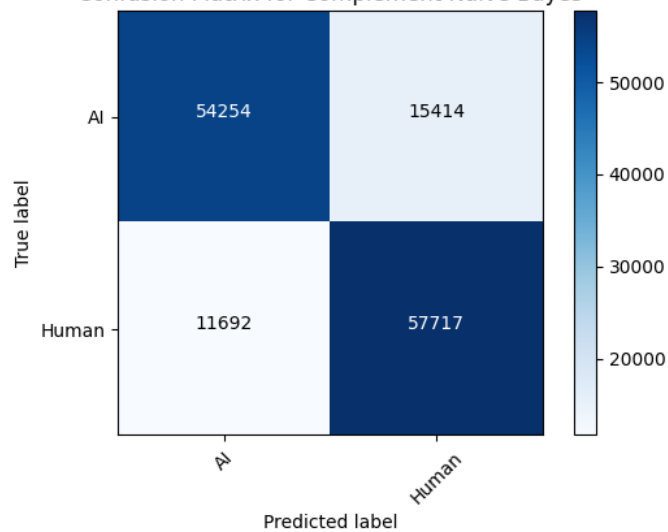
6.3 ROC modello C-NB

Receiver Operating Characteristic (ROC) Curve for Complement Naive Bayes



6.4 Confusion Matrix modello C-NB

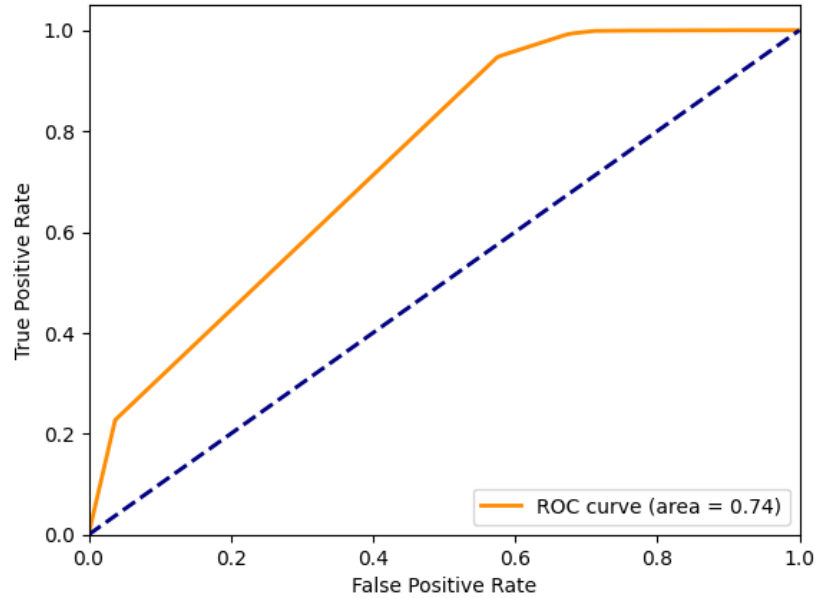
Confusion Matrix for Complement Naive Bayes





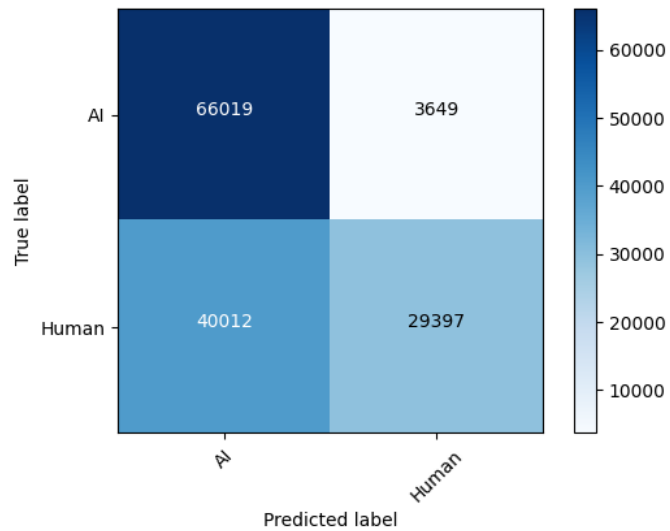
6.5 ROC modello DTC

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier



6.6 Confusion Matrix modello DTC

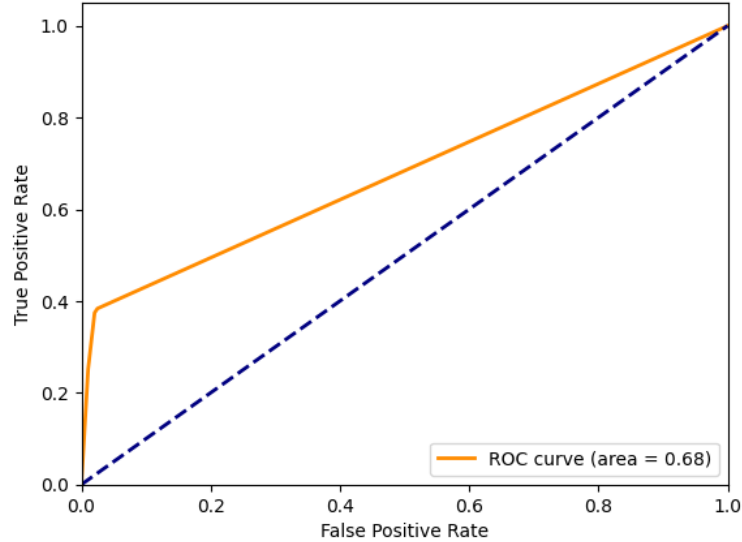
Confusion Matrix for Decision Tree Classifier



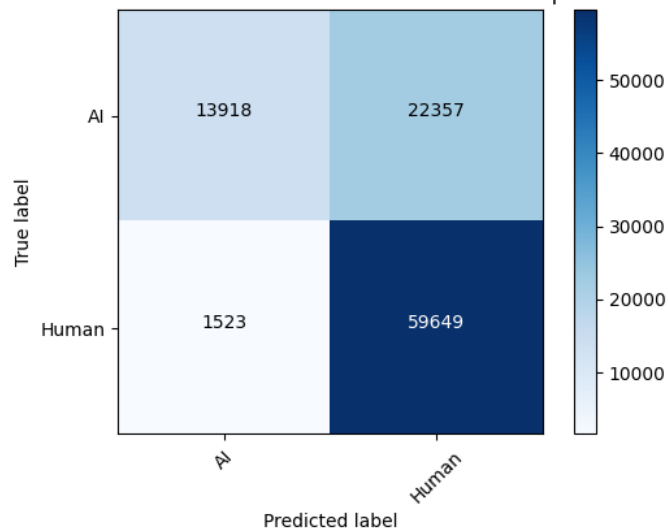
7. DTC varie profondità – Dataset 1 non bilanciato

7.1 Profondità 2

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=2



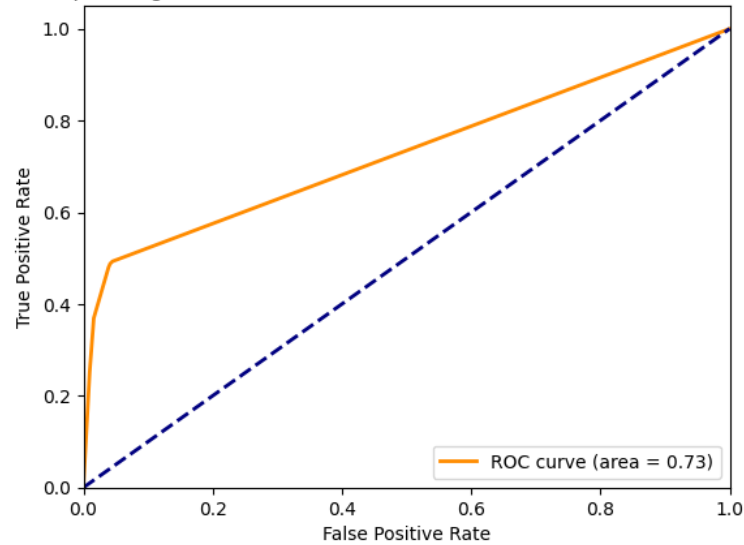
Confusion Matrix for Decision Tree Classifier depth=2



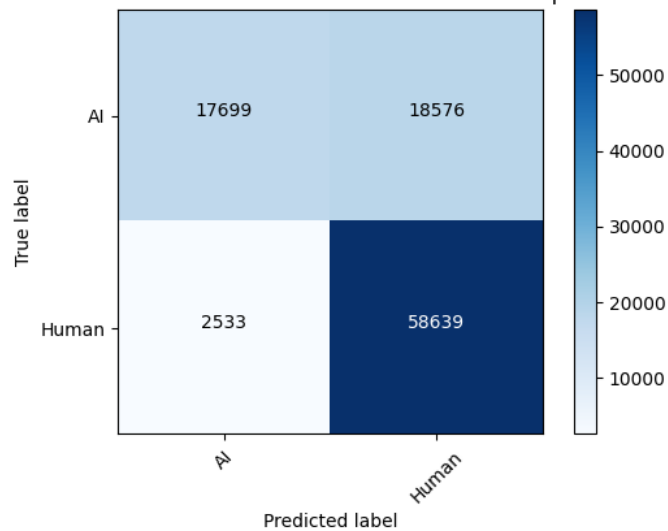


7.2 Profondità 3

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=3



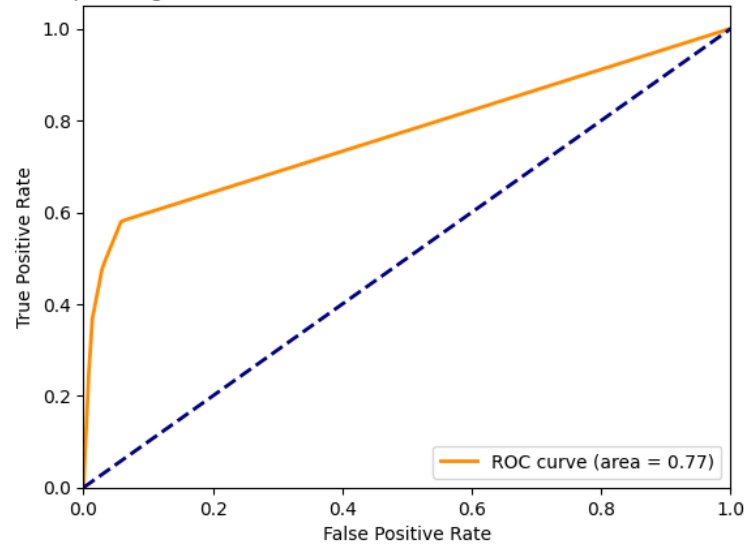
Confusion Matrix for Decision Tree Classifier depth=3



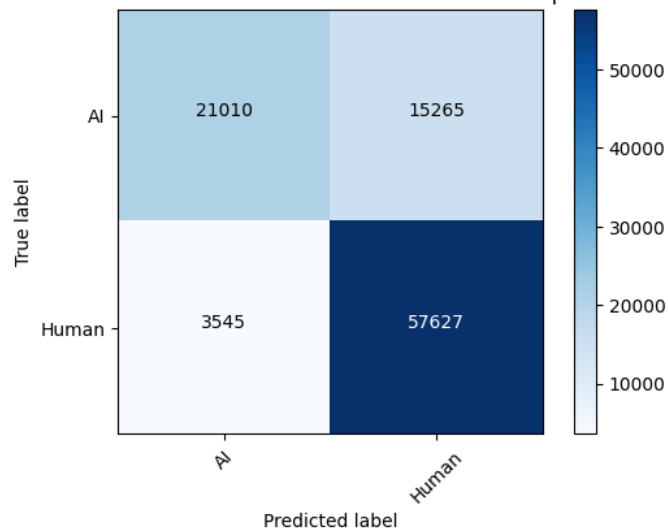


7.3 Profondità 4

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=4



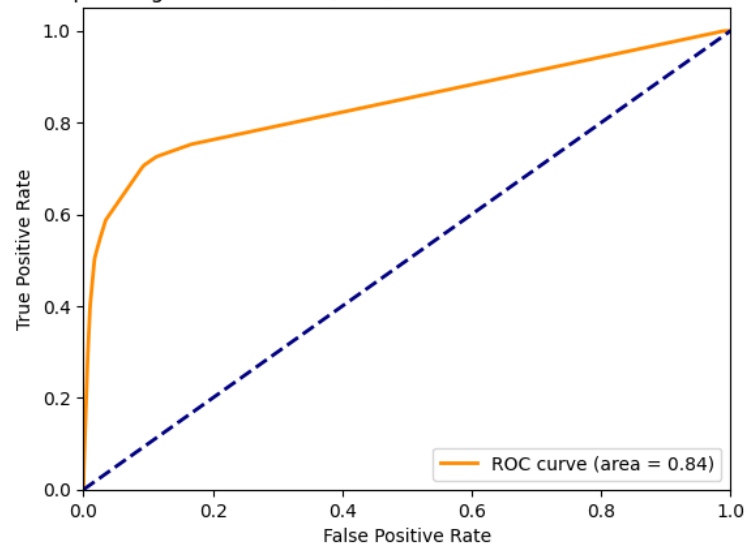
Confusion Matrix for Decision Tree Classifier depth=4



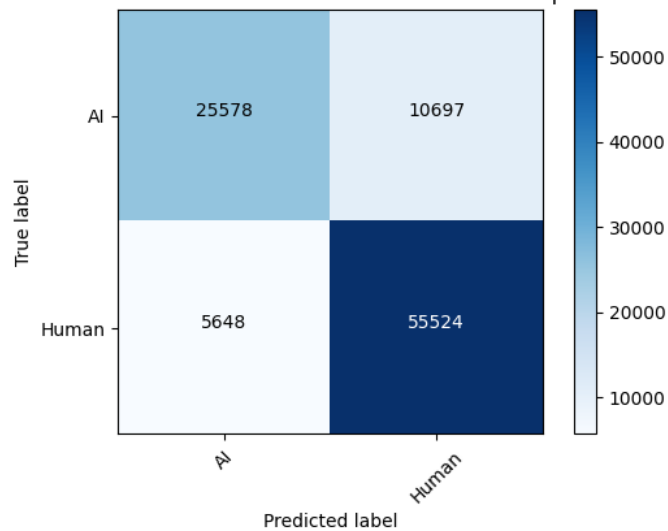


7.4 Profondità 6

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=6



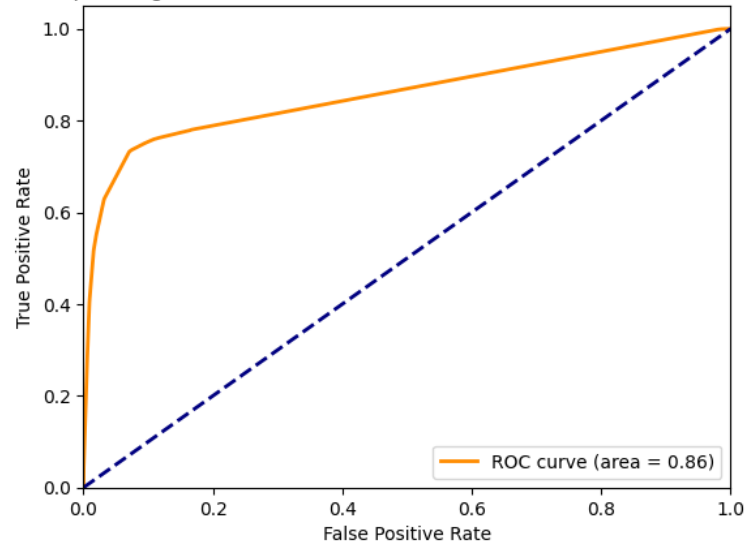
Confusion Matrix for Decision Tree Classifier depth=6



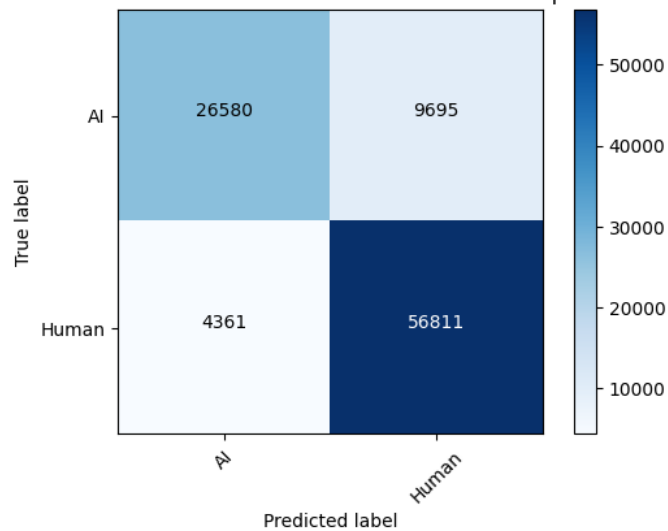


7.5 Profondità 7

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=7



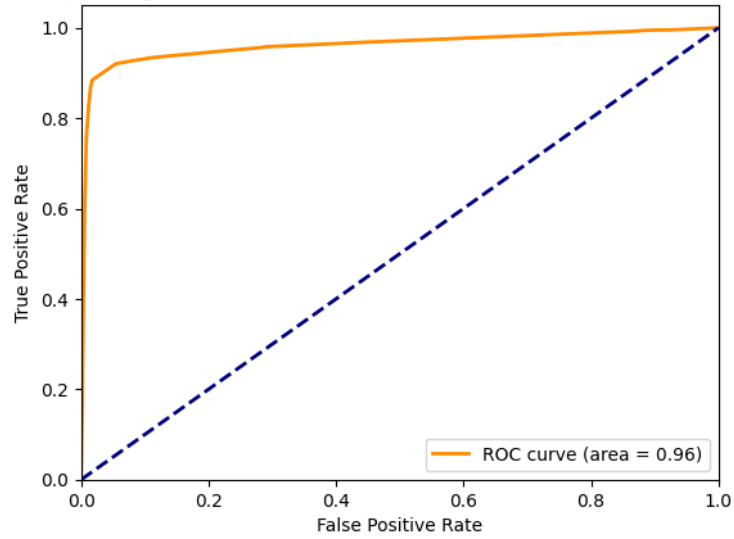
Confusion Matrix for Decision Tree Classifier depth=7



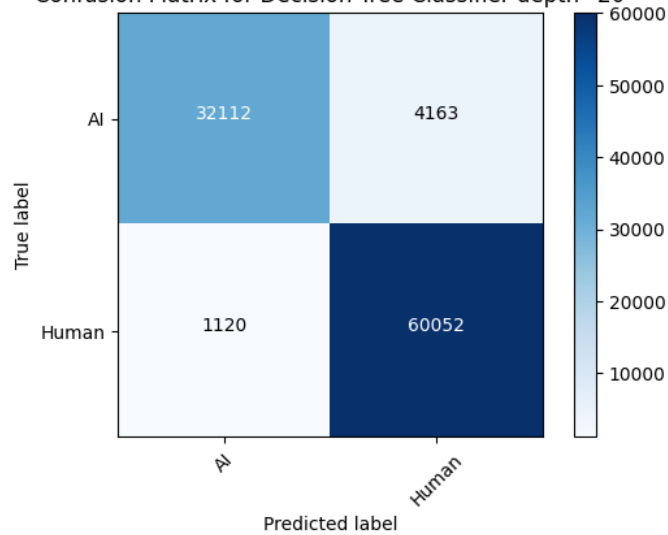


7.6 Profondità 20

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=20



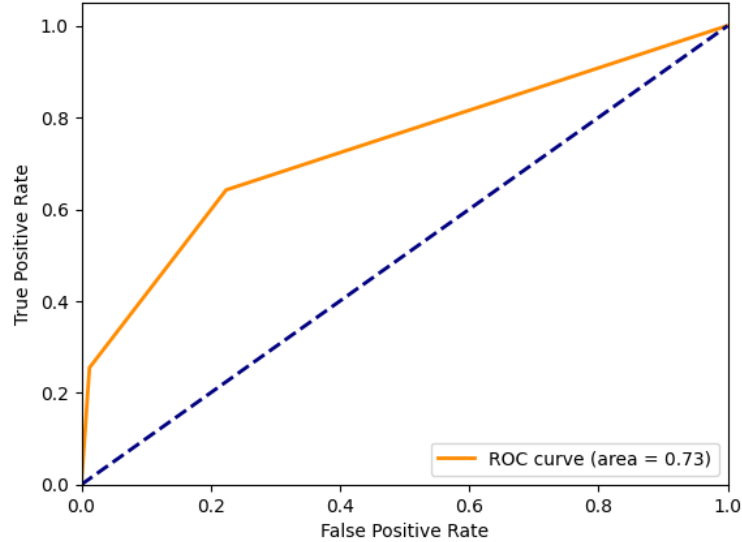
Confusion Matrix for Decision Tree Classifier depth=20



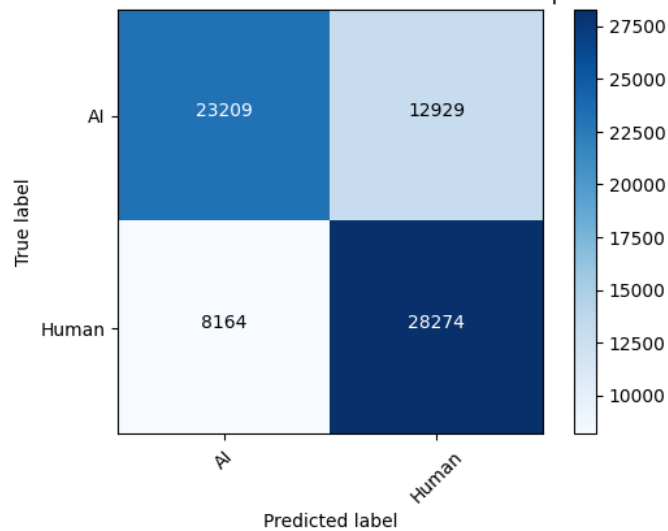
8. DTC varie profondità – Dataset 1 bilanciato

8.1 Profondità 2

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=2



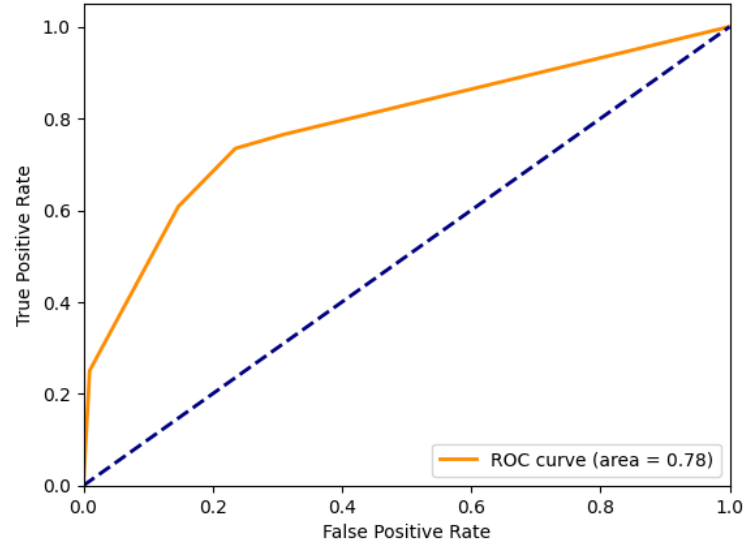
Confusion Matrix for Decision Tree Classifier depth=2



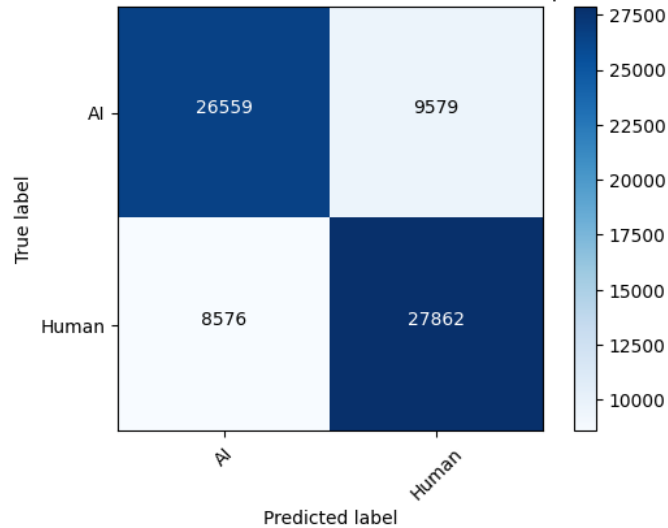


8.2 Profondità 3

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=3



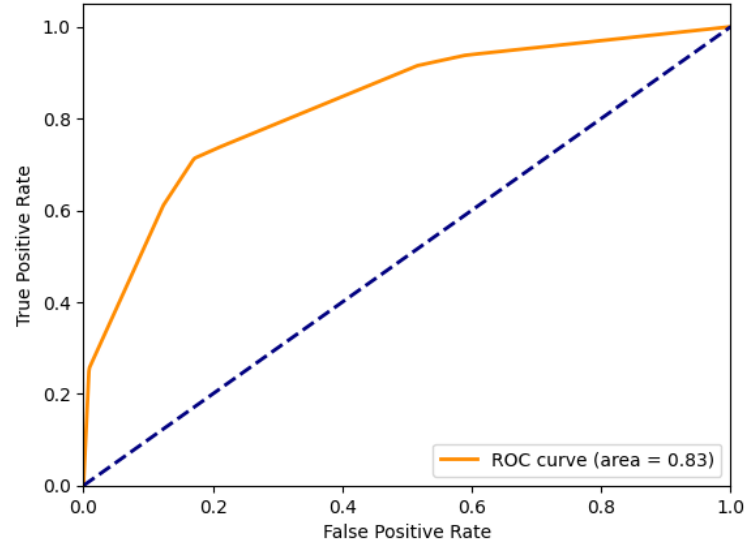
Confusion Matrix for Decision Tree Classifier depth=3



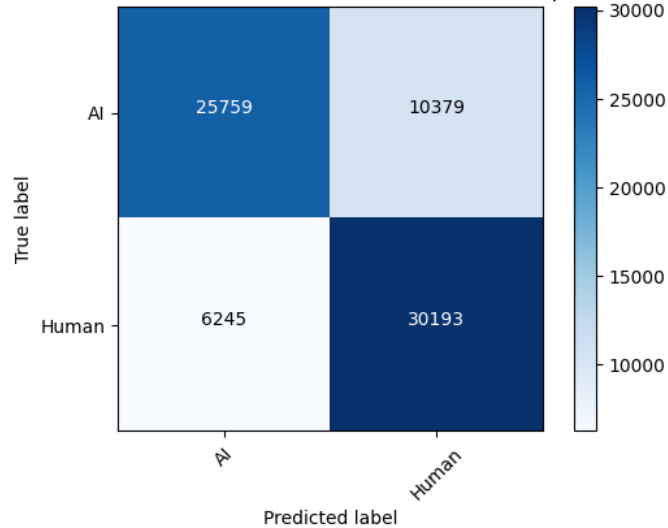


8.3 Profondità 4

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=4



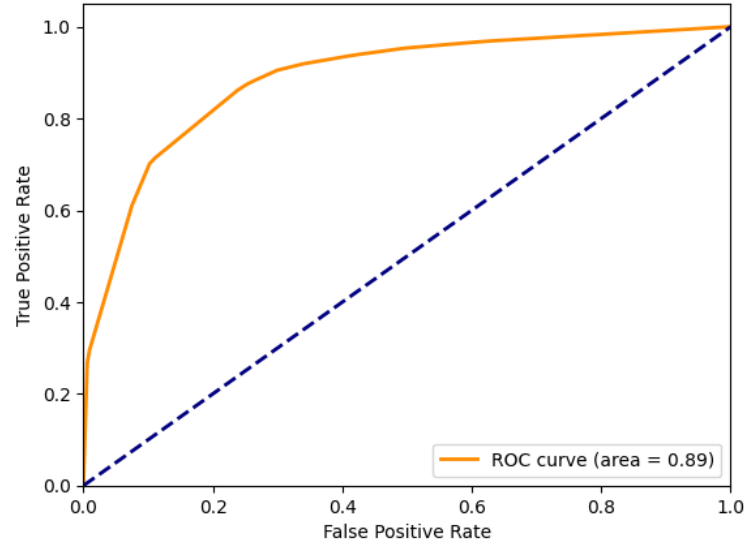
Confusion Matrix for Decision Tree Classifier depth=4



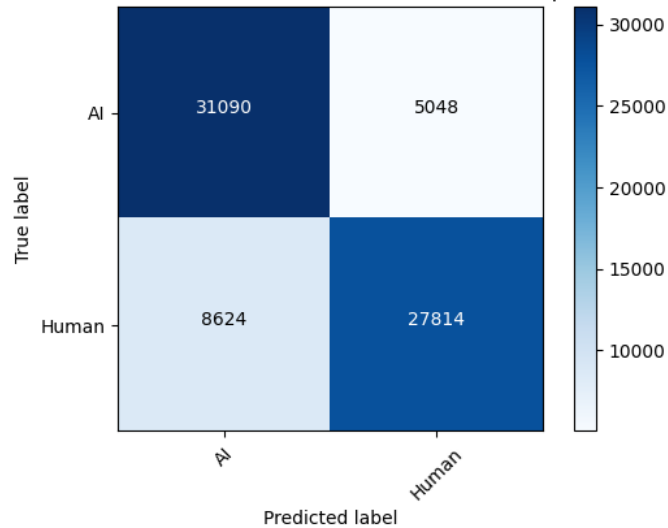


8.4 Profondità 6

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=6



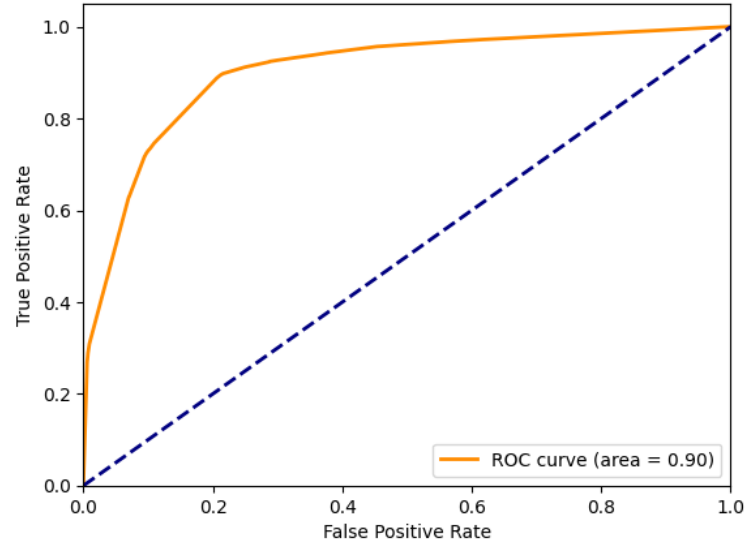
Confusion Matrix for Decision Tree Classifier depth=6



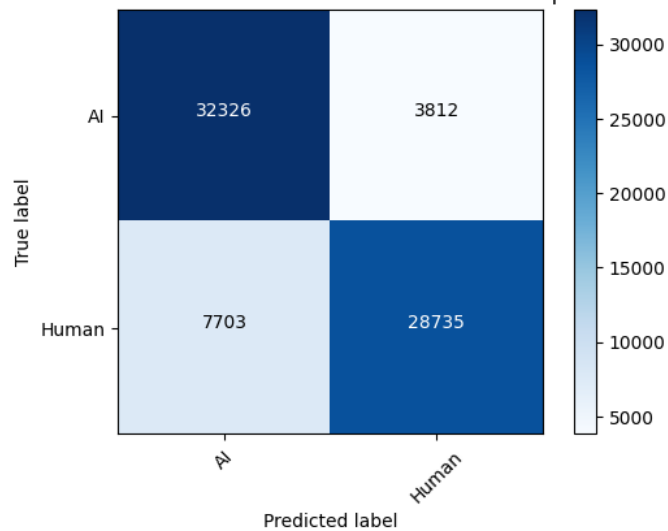


8.5 Profondità 7

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=7



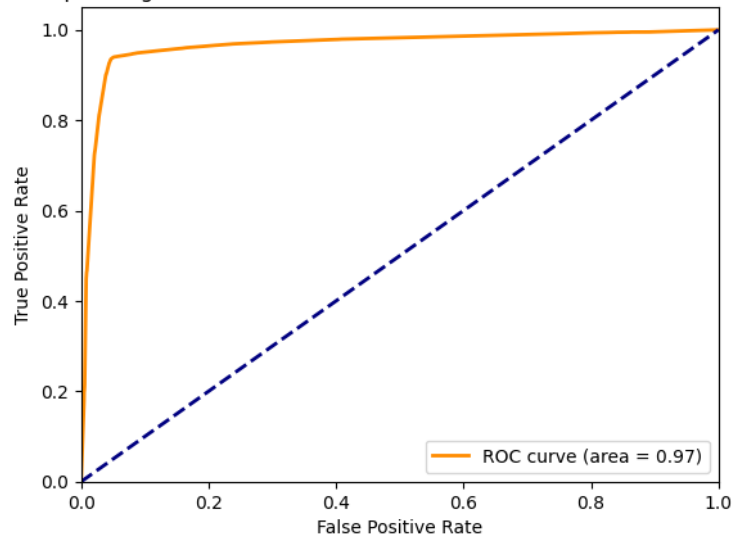
Confusion Matrix for Decision Tree Classifier depth=7



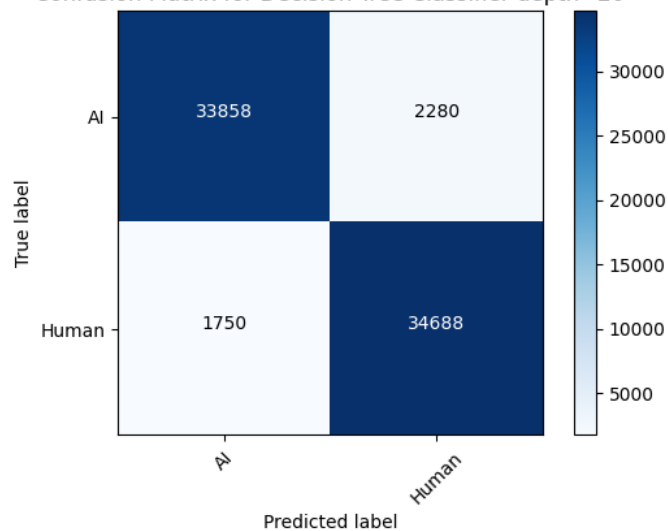


8.6 Profondità 20

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=20



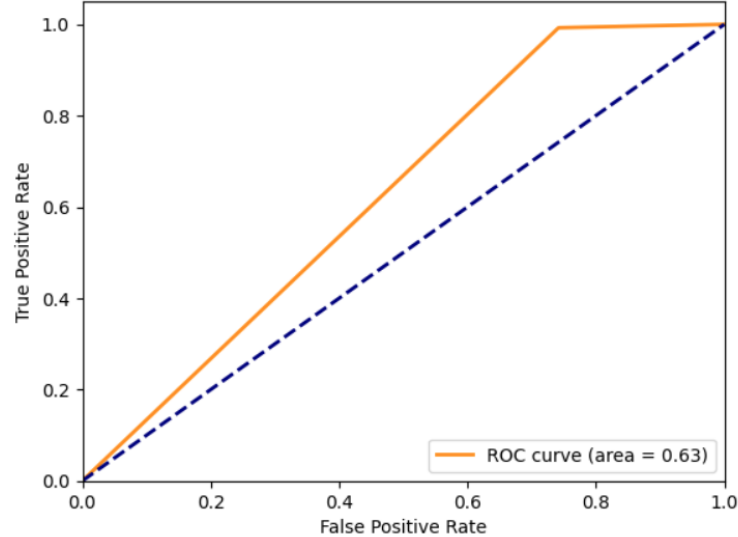
Confusion Matrix for Decision Tree Classifier depth=20



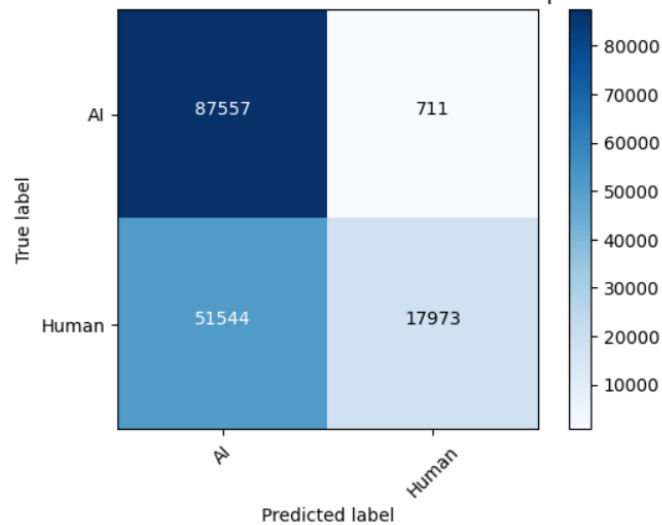
9. DTC varie profondità – Dataset 2 non bilanciato

9.1 Profondità 2

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=2



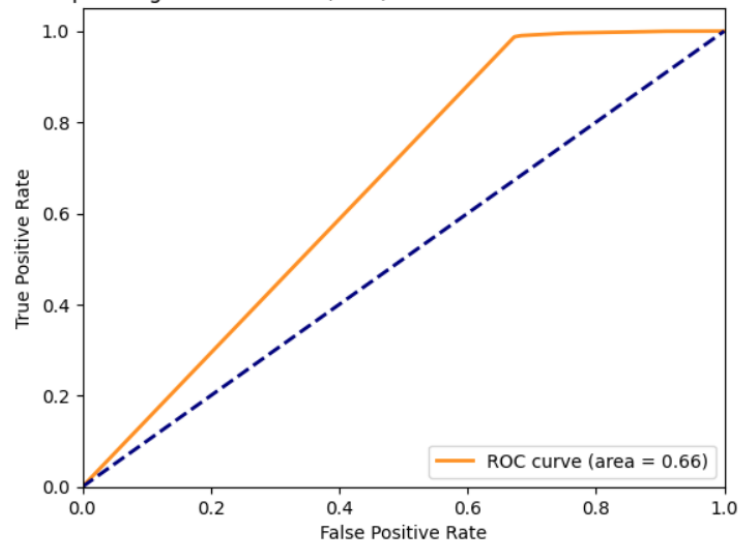
Confusion Matrix for Decision Tree Classifier depth=2



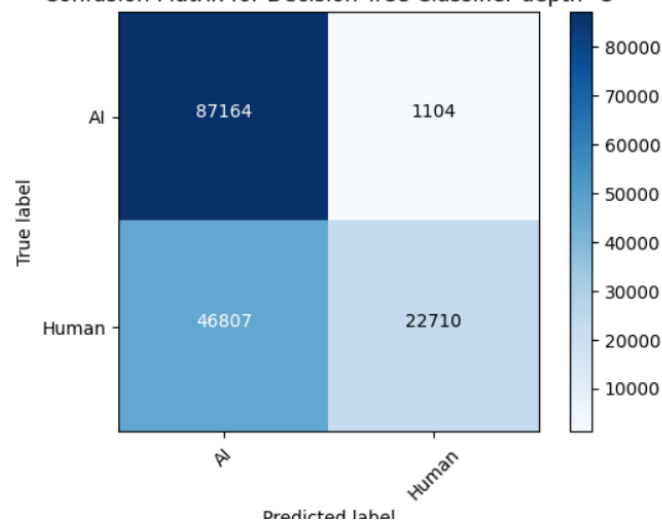


9.2 Profondità 3

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=3



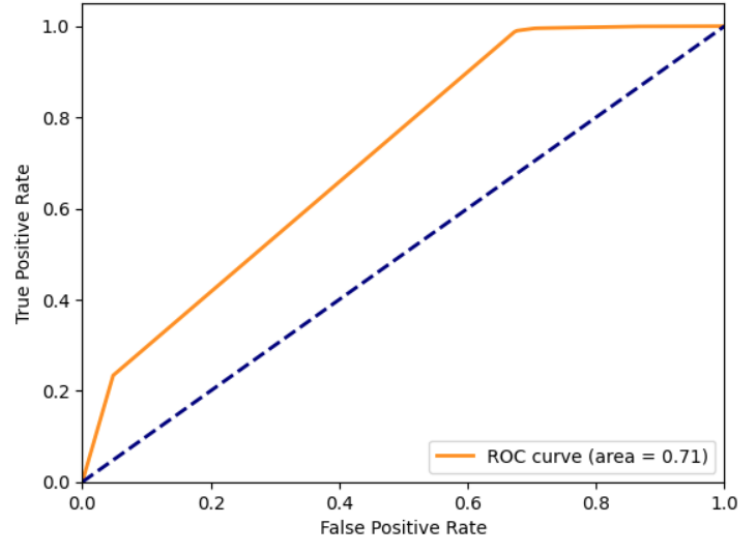
Confusion Matrix for Decision Tree Classifier depth=3



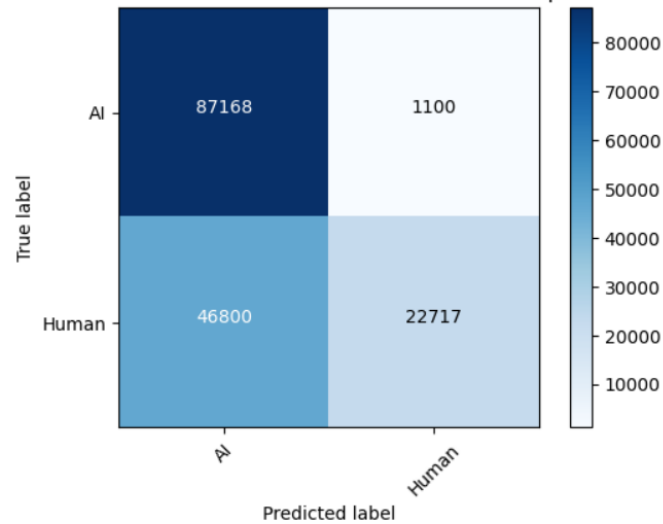


9.3 Profondità 4

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=4



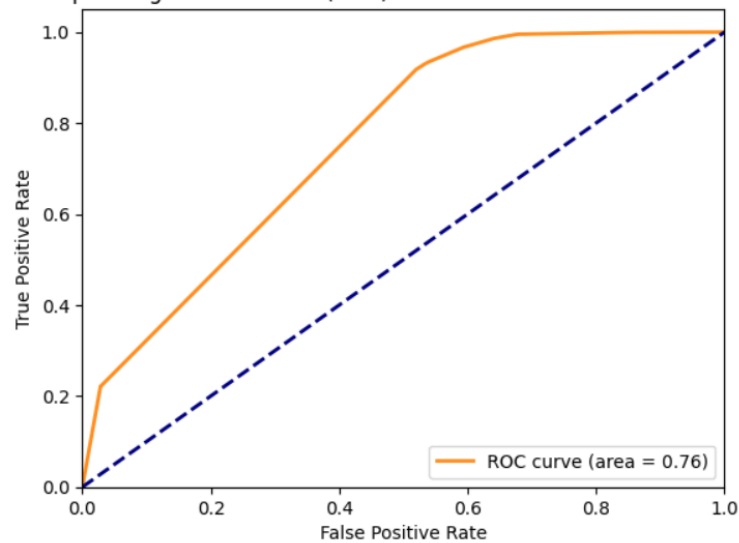
Confusion Matrix for Decision Tree Classifier depth=4



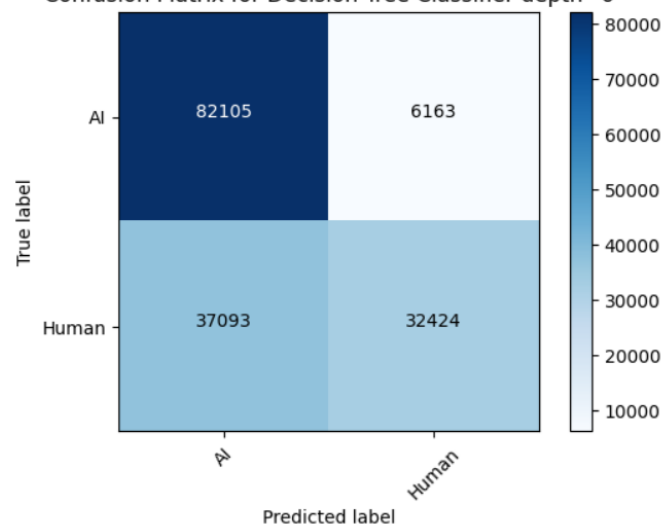


9.4 Profondità 6

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=6



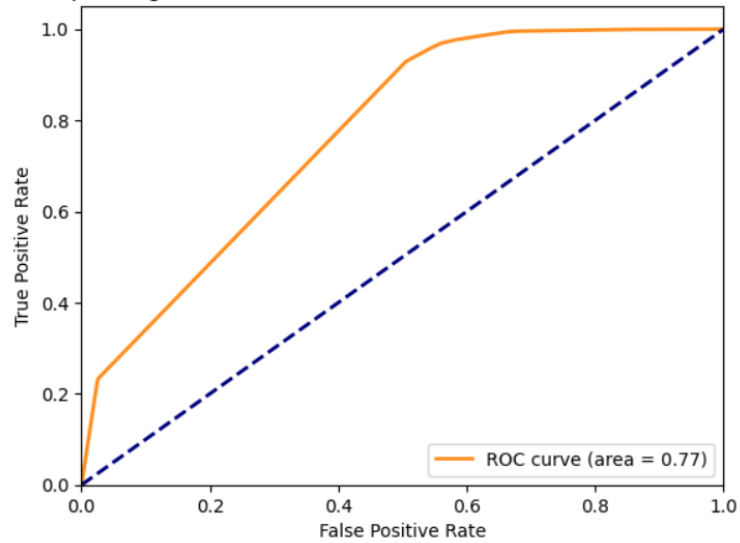
Confusion Matrix for Decision Tree Classifier depth=6



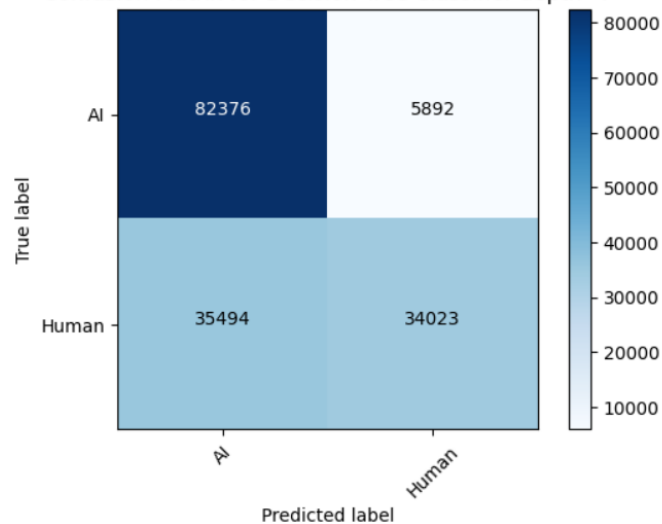


9.5 Profondità 7

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=7



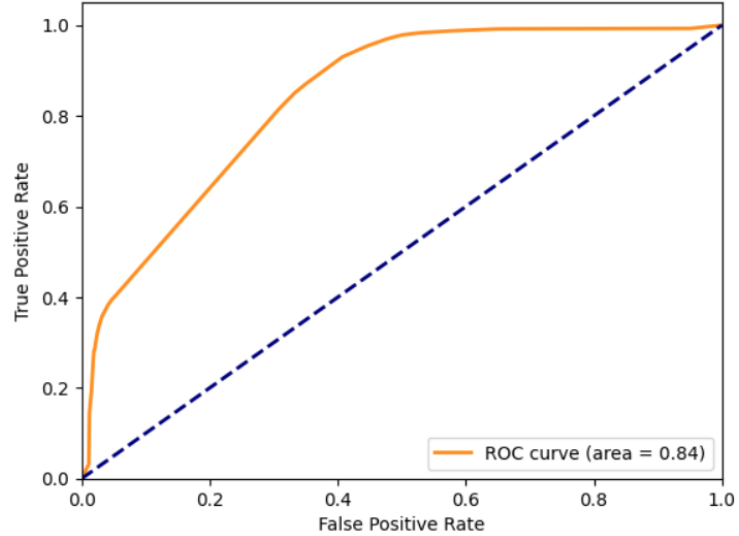
Confusion Matrix for Decision Tree Classifier depth=7



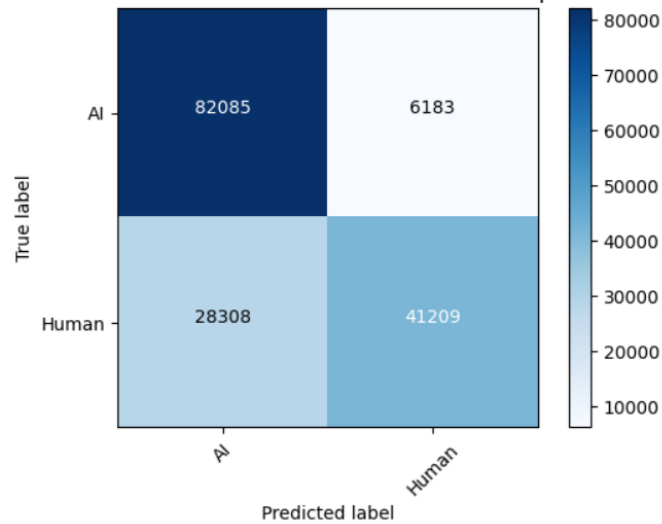


9.6 Profondità 20

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=20



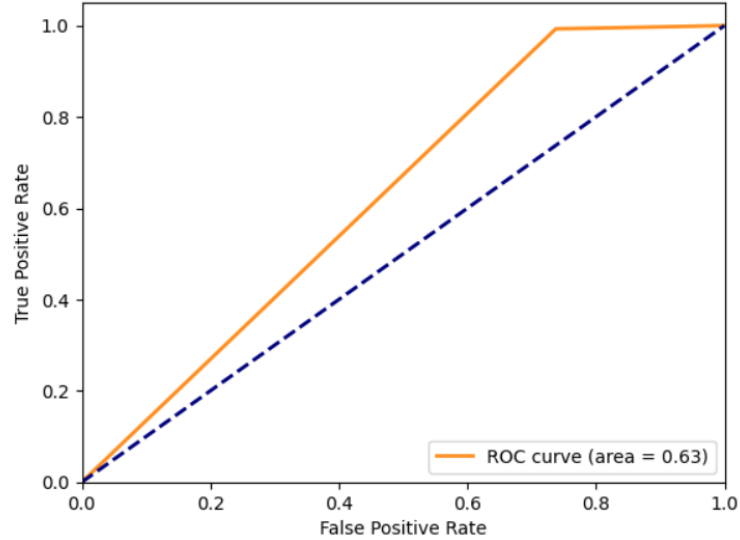
Confusion Matrix for Decision Tree Classifier depth=20



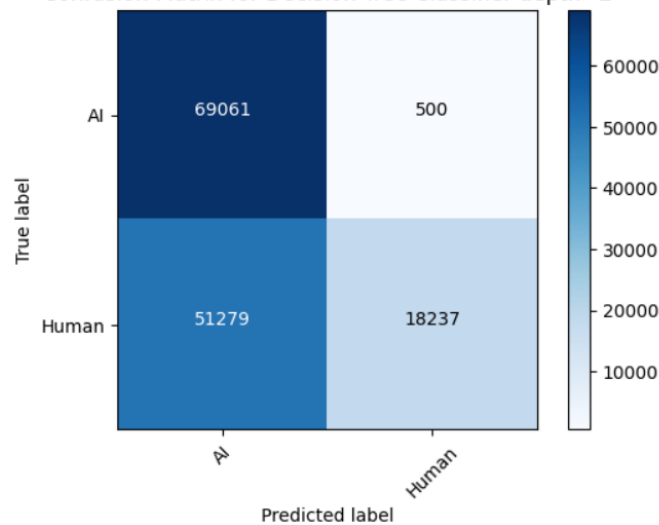
10. DTC varie profondità – Dataset 2 bilanciato

10.1 Profondità 2

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=2



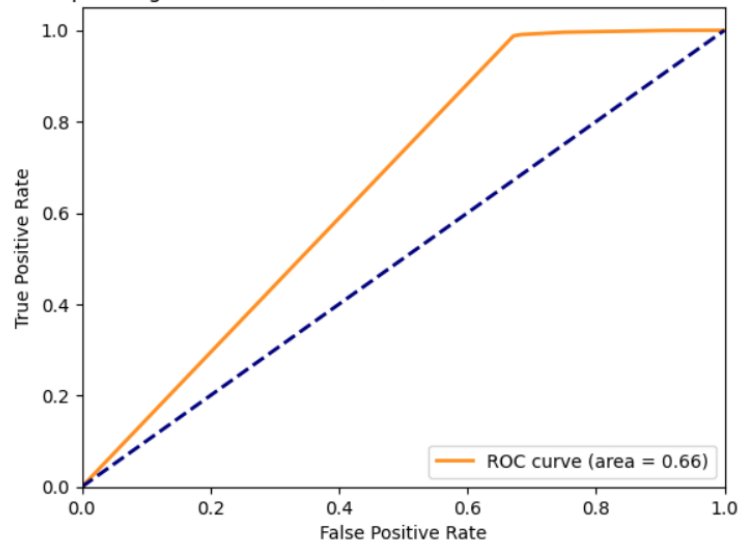
Confusion Matrix for Decision Tree Classifier depth=2



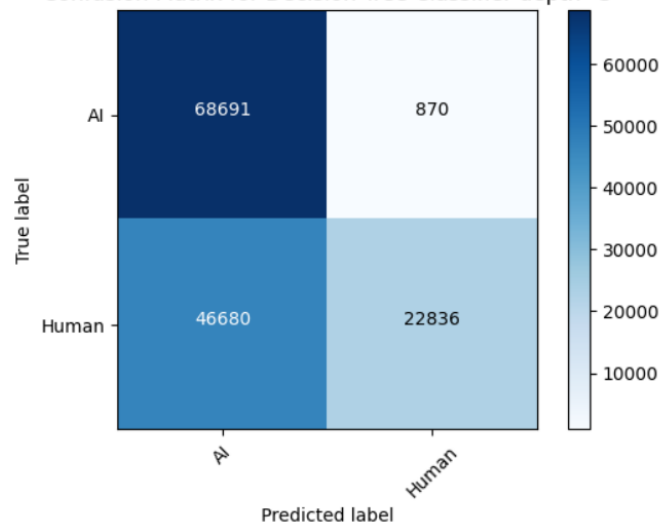


10.2 Profondità 3

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=3



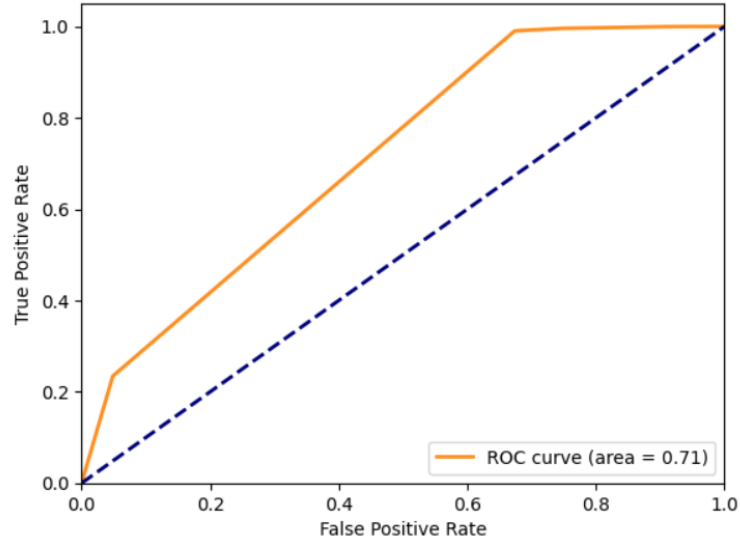
Confusion Matrix for Decision Tree Classifier depth=3



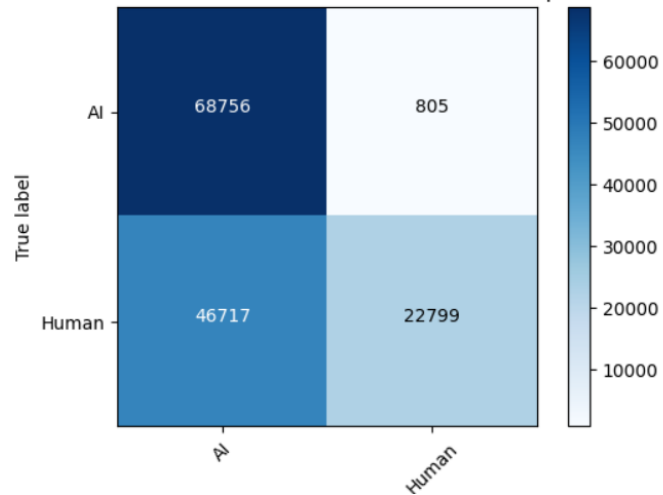


10.3 Profondità 4

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=4



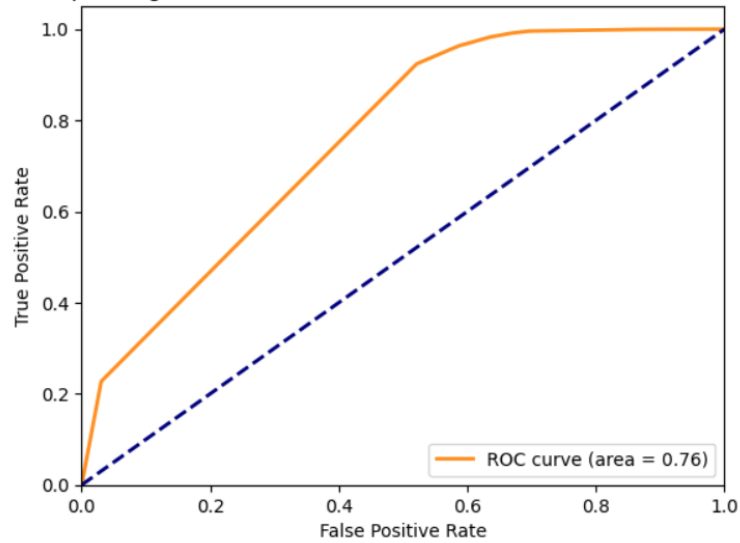
Confusion Matrix for Decision Tree Classifier depth=4



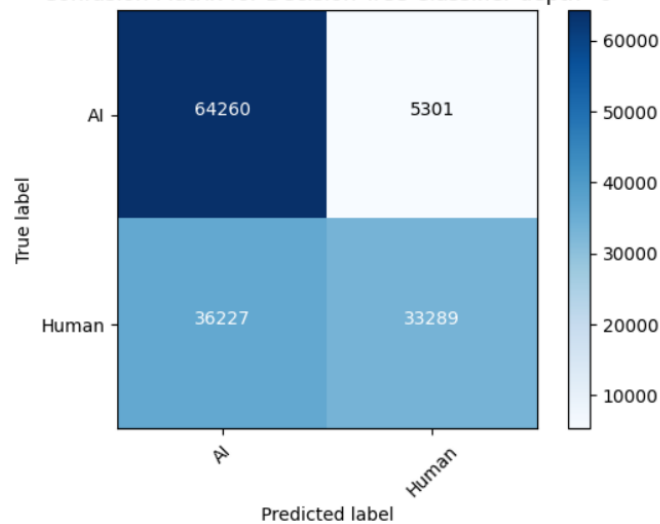


10.4 Profondità 6

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=6



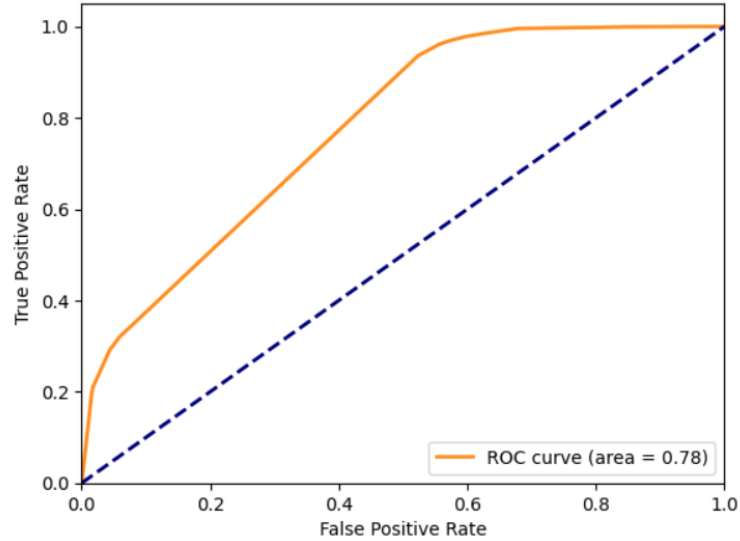
Confusion Matrix for Decision Tree Classifier depth=6



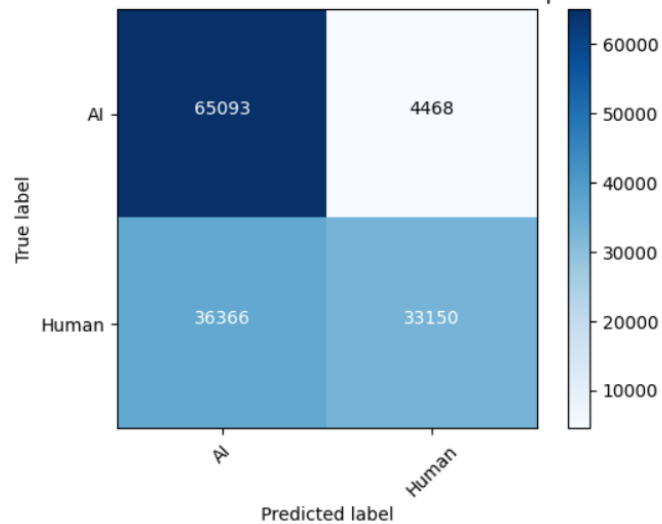


10.5 Profondità 7

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=7



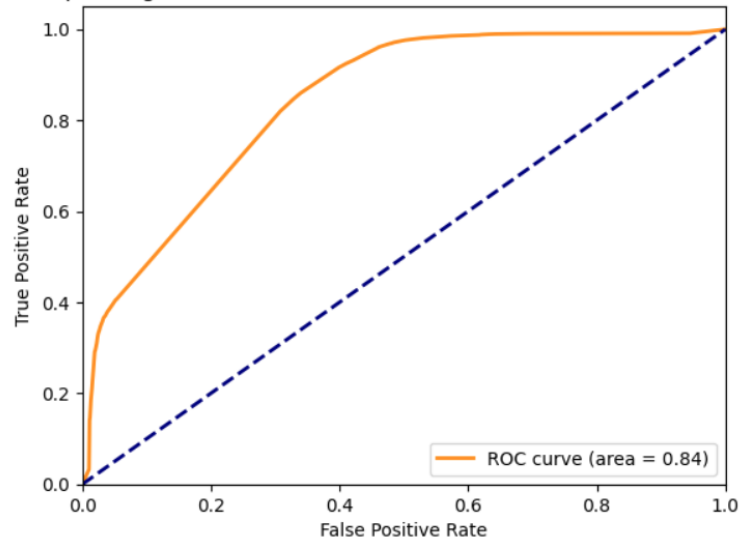
Confusion Matrix for Decision Tree Classifier depth=7





10.6 Profondità 20

Receiver Operating Characteristic (ROC) Curve for Decision Tree Classifier depth=20



Confusion Matrix for Decision Tree Classifier depth=20

