5. חווה דעתך עד כמה תוצאות המודל קרובות ל Training:

AUC on Validation File: 0.999244142101285

AUC on Training: 0.9998437011566115

הערך של AUCשל validation קרוב לשל training שקרוב ל1 , וזה מצביע על כך שלמודל יש כוח חיזוי גבוה מאוד והוא מפריד למעשה בין המחלקות החיוביות והשליליות .

confusion matrix on validation: [[13677 0]

[ 2 1321]]

confusion matrix on Training: [[31801 0]

[ 1 3198]]

שניהם מראים ביצועים גבוהים, מספר הFN והFP נמוך מאוד בהשוואה ל TN ו TP, מה שמצביע על יכולת חיזוי חזקה של המודל. בנוסף, הערכים הגבוהים של TP ו TN מרמזים על כך שהמודל מסווג ביעילות את המקרים גם ב validation וגם ב training data.

Accuracy train: 0.9999714285714286

Precision train: 1.0

Recall train: 0.9996874023132228

F1-Score train: 0.9998436767234641

Accuracy validation: 0.9998666666666667

Precision validation: 1.0

Recall validation: 0.99848828420257

F1-Score validation: 0.9992435703479576

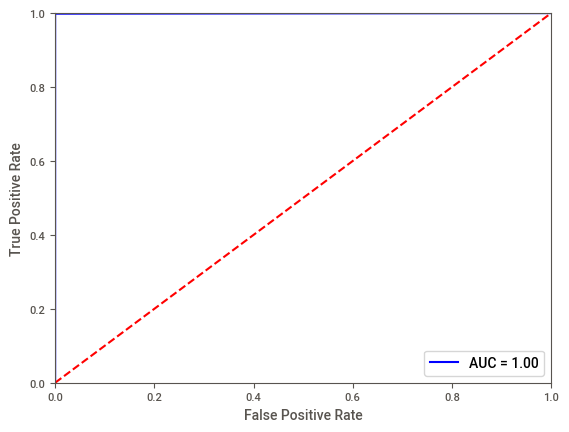
ניתן לראות , שערכי המדדים האלה עבור validation ממש קרובים לאלה של training . בנוסף , מדדים אלה מראים שהמודל מתפקד טוב מאוד גם בvalidation וגם ב training data, עם ערכי דיוק גבוהים .

expected value validation : 132100

expected value Training : 319800

הערך של הtraining יותר גבוה , שמעיד שהביצוע על הנתונים האלה יותר טוב .

ROC curves for validation :



For training :

