

למידת מכונה – פרוייקט גמר: סיווג קולות של חיות

מגישות:

שיר אברהם

נועה יאיר

קוד:

<https://github.com/Shir-Av/MachineLearning-FinalProject.git>

תיאור המאגר:

המאגר שלנו מכיל קבצי קול (WAV) של שני סוגים של חיות - חיות ים וציפורים.

מאגר חיות הים מכיל 1682 הקלטות של 30 זנים שונים.

מאגר הציפורים מכיל 2000 הקלטות של 20 זנים שונים.

המאגר מסודר בתיקיות בהיררכיה לפי סוג החיה (חיות ים או ציפורים) ומחולק לתיקיות לפי זנים.

את קבצי הקול אנו ממירים לווקטורים באמצעות ספרייה בPython בשם librosa, הספרייה מאפשרת לטעון קובץ שמע ולחלץ ממנו מידע אודות הקובץ, היא עושה זאת באמצעות הוצאת פיצ'רים בדידים מקטעי הקול ע"י דגימה של האות וביצוע מניפולציות מתמטיות מתחום עיבוד האותות, לבסוף מחזירה ווקטור המייצג את קובץ השמע.

בעיות שפתרנו:

1. סיווג בין חיות ים לציפורים –

השאלה הראשונה שניסינו לענות עליה היא האם ניתן לסווג בין חיות ים לציפורים באחוזים גבוהים, רצינו לבדוק האם יש הבדל משמעותי בין קולות חיות הים לקולות הציפורים, או שיש גורם משפיע נוסף על הסיווג, קיבלנו תוצאות טובות בדיוק מרשים של 95%-99%.
אנחנו משערות שמעבר לקולות השונים עצם פעולת ההקלטה מתחת למים מבדילה את קטעי הקול של חיות הים מחיות היבשה (ציפורים).
אם אכן הדבר קיים, נוכל להסיק שעצם ההקלטה מתחת למים נותנת תכונה ייחודית המאפשרת סיווג בדיוק גבוה ומוכיח כי ניתן לדעת האם קטע האודיו הנ"ל הוקלט מתחת למים או לא, בלי קשר לחיה שהוקלטה.

2. סיווג בין חיות לפי זן:

שאלה זו מתמקדת בסיווג סוג חיה לפי זנים (סיווג ציפורים לפי זני ציפורים וחיות ים לפי זני חיות ים), לקחנו את שני המאגרים (ציפורים וחיות ים) וסיווגנו כל אחד בנפרד לפי זנים.
רצינו לבדוק איזה סוג חיה ייתן לנו תוצאות טובות יותר ולהסיק מכך על ההבדלים בקולות החיות בשני הסוגים.
במאגר חיות הים קיבלנו תוצאות טובות של 88%-92%.
במאגר הציפורים קיבלנו תוצאות לא כל כך טובות בדיוק נמוך של 39%-49%.
(טווח הדיוק גדול מכיוון שהצלחנו לשפר באמצעות הגדלת הווקטור, יפורט בהמשך)

3. עבודה עם וקטורים במימדים שונים:

שאלה זו עוסקת בהשפעה של עבודה עם וקטורים במימדים שונים על הסיווגים השונים. רצינו לבדוק האם שינוי מימד (גודל) וקטור הנתונים ישפיע על סיווג המאגרים, ואם כן, כיצד?

יצרנו 2 סוגים של וקטורים, וקטור אחד עם 13 נתונים ווקטור עם 25 נתונים. ביצענו כל אחד מהסיווגים (שהזכרנו לעיל בבעיות 1 ו-2) עם שני הווקטורים השונים.

בסיווג החיות לפי סוג חיה ובסיווג מאגר חיות הים ראינו שיפור קטן של בערך 2%-5%, ובמאגר הציפורים ראינו שיפור קטן עוד יותר של 1%-2%.

4. סיווג קטעי קול בהם נשמעים 2 זנים שונים:

שאלה זו עוסקת בעובדה שיתכן שזנים שונים של חיות ימצאו באזורים מסויימים באותו זמן ורצינו לבדוק אם נוכל לקבל תוצאות טובות וסיווג נכון גם כאשר נשמע אותם יחד. לאחר שקיבלנו את התוצאות של סיווגי זני חיות הים וזני הציפורים ומצאנו כי סיווג זני הציפורים לא הניב תוצאות טובות ניסינו לסווג זוגות של חיות ים שקיבלנו עליהם את התוצאות הכי טובות בסיווג. לקחנו 2 הקלטות בעלות אורך זהה בהן נשמעות החיות בצורה ברורה, ובאמצעות תוכנה לעיבוד קבצי אודיו מיזגנו אותם יחד שינגנו באותו הזמן, הסיווגים לא הניבו תוצאות כל כך טובות, לרוב רק חיה אחת סווגה נכון וקיבלנו סיווג שגוי על החיה השנייה. שיערנו שזה נובע מהשילוב שיצרנו, ואולי אם הייתה לנו הקלטה משולבת מקורית היינו מקבלות תוצאות טובות יותר.

טכניקות בהן השתמשנו:

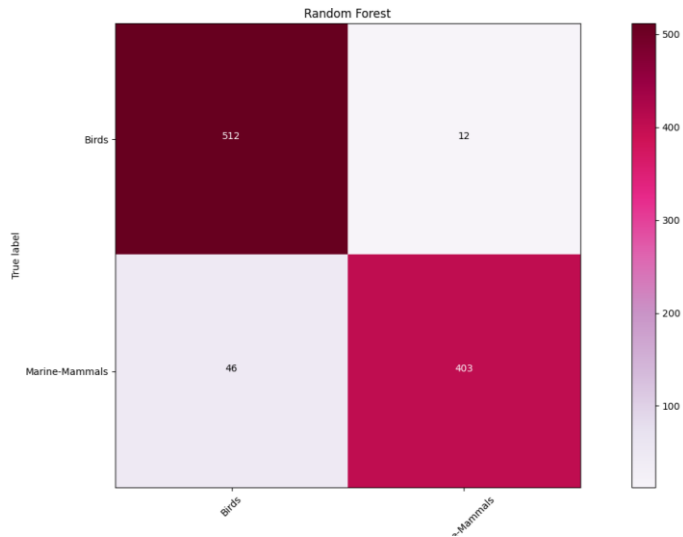
1. אלגוריתם Nearest Neighbor
2. אלגוריתם SVM
3. אלגוריתם Random Forest
4. שימוש ב PCA על מנת לשחק עם גודל/ מימד הווקטורים שיצרנו.

תוצאות הסיווגים:

1. סיווג בין חיות ים לציפורים –

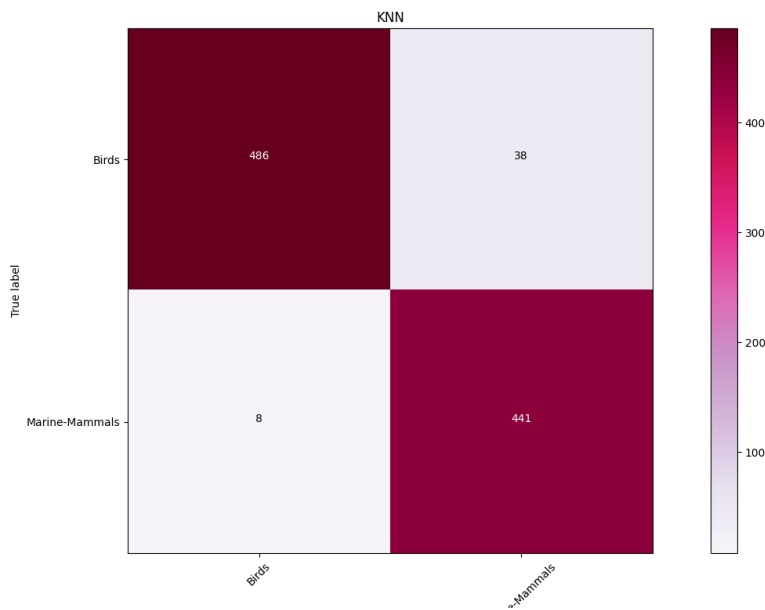
• וקטור במימד 13 :

Random Forest:



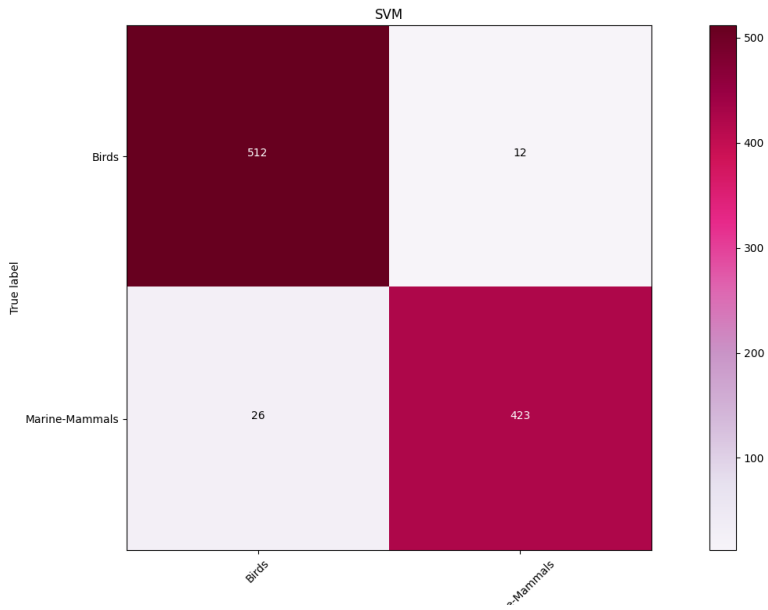
```
----- Random Forest - Results -----
Recall:  [0.98 0.9 ]
Precision: [0.92 0.97]
F1-Score: [0.95 0.93]
Accuracy: 0.94 , 915
Number of samples: 973
```

K-Nearest Neighbor:



```
----- KNN - Results -----
Recall:  [0.92748092 0.98218263]
Precision: [0.98380567 0.92066806]
F1-Score: [0.95481336 0.95043103]
Accuracy: 0.95 , 927
Number of samples: 973
```

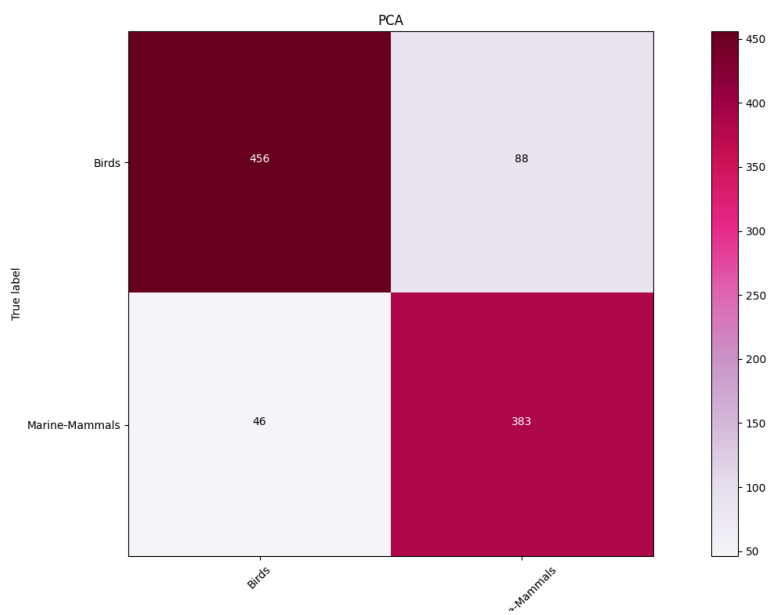
SVM:



----- SVM - Results

```
Recall:  [0.98 0.94]
Precision: [0.95 0.97]
F1-Score: [0.96 0.96]
Accuracy: 0.96 , 935
Number of samples: 973
```

PCA:

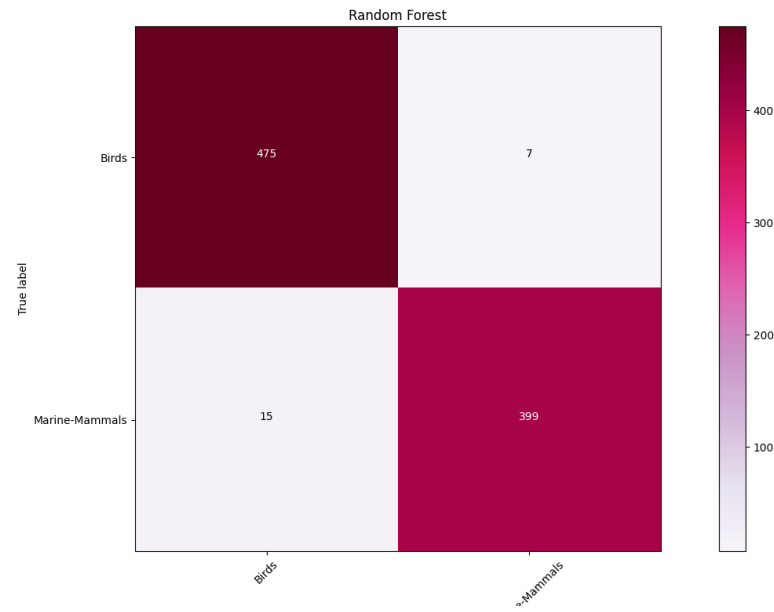


----- PCA - Results

```
Recall:  [0.83823529 0.89277389]
Precision: [0.90836653 0.81316348]
F1-Score: [0.87189293 0.85111111]
Accuracy: 0.86 , 839
Number of samples: 973
```

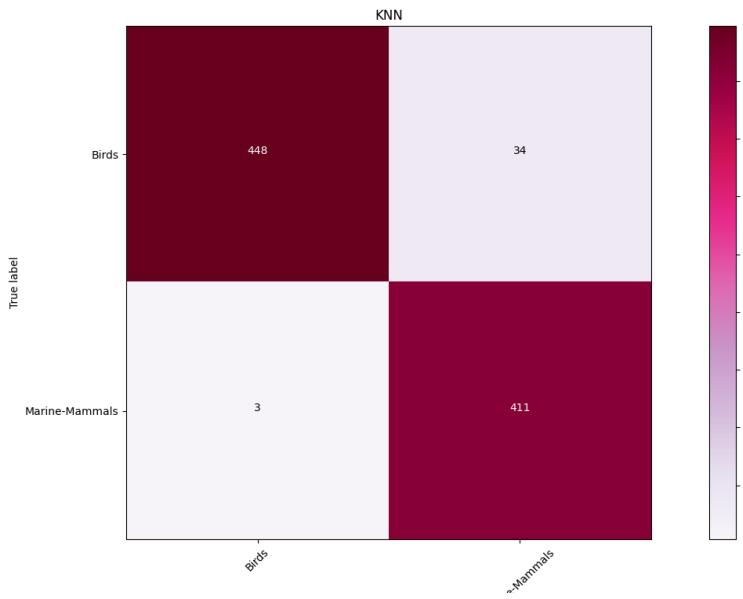
• וקטור במימד 25 :

Random Forest:



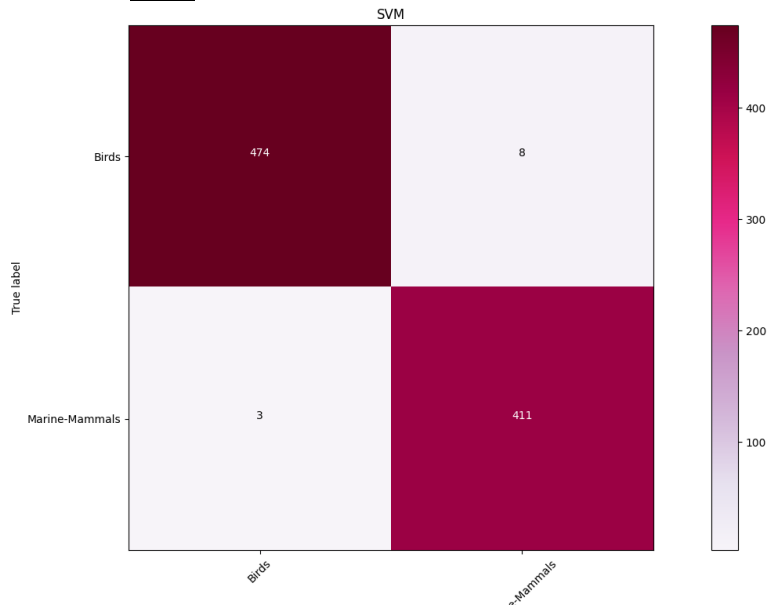
```
----- Random Forest - Results
Recall:  [0.99 0.96]
Precision: [0.97 0.98]
F1-Score: [0.98 0.97]
Accuracy: 0.98 , 874
Number of samples: 896
```

K-Nearest Neighbor:



```
----- KNN - Results
Recall:  [0.92946058 0.99275362]
Precision: [0.99334812 0.92359551]
F1-Score: [0.96034298 0.95692666]
Accuracy: 0.96 , 859
Number of samples: 896
```

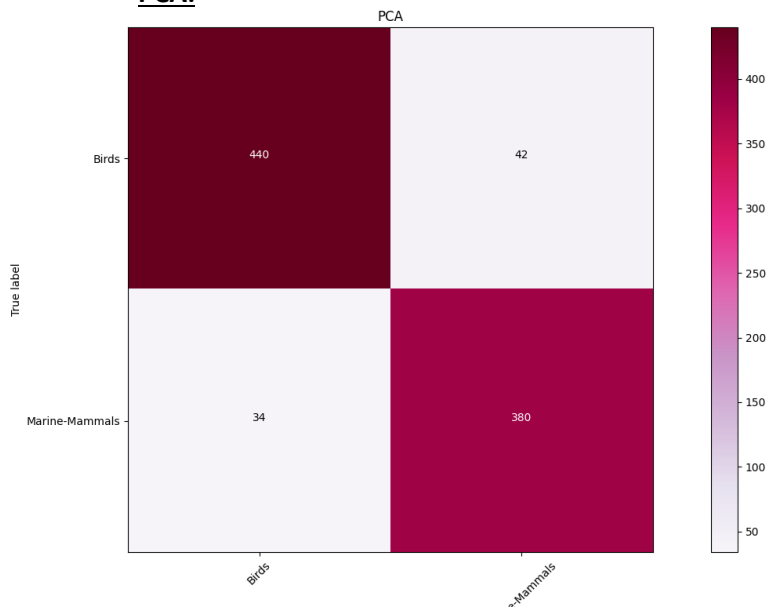
SVM:



----- SVM - Results

```
Recall:  [0.98 0.99]
Precision: [0.99 0.98]
F1-Score: [0.99 0.99]
Accuracy: 0.99 , 885
Number of samples: 896
```

PCA:



----- PCA - Results

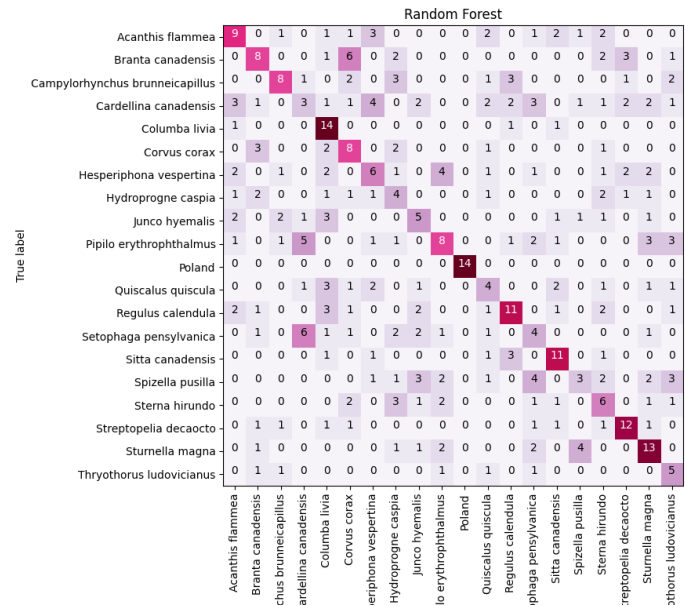
```
Recall:  [0.91 0.92]
Precision: [0.93 0.9 ]
F1-Score: [0.92 0.91]
Accuracy: 0.92 , 820
Number of samples: 896
```

2. סיווג בין חיות לפי זן –

סיווג ציפורים:

- וקטור במימד 13 :

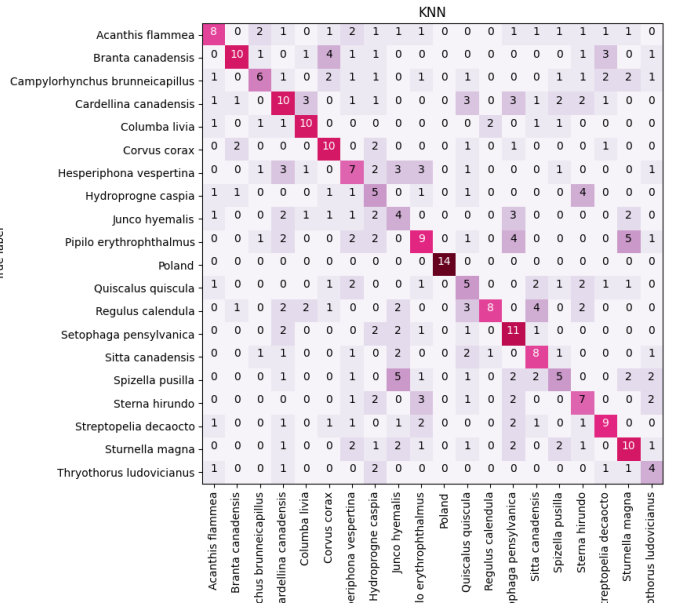
Random Forest:



```
----- Random Forest - Results -----
Recall:  [0.39 0.35 0.38 0.1 0.82 0.47 0.26 0.27 0.29 0.3 1. 0.24 0.44 0.2
0.61 0.14 0.33 0.6 0.54 0.5 ]
Precision: [0.43 0.42 0.53 0.18 0.41 0.32 0.32 0.2 0.29 0.4 1. 0.24 0.52 0.2
0.52 0.3 0.26 0.57 0.46 0.28]
F1-Score: [0.41 0.38 0.44 0.13 0.55 0.38 0.29 0.23 0.29 0.34 1. 0.24 0.48 0.2
0.56 0.19 0.29 0.59 0.5 0.36]
Accuracy: 0.39 , 156
Number of samples: 400
```



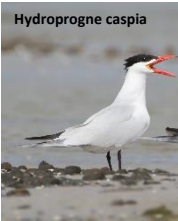
K-Neareast Neighbor:



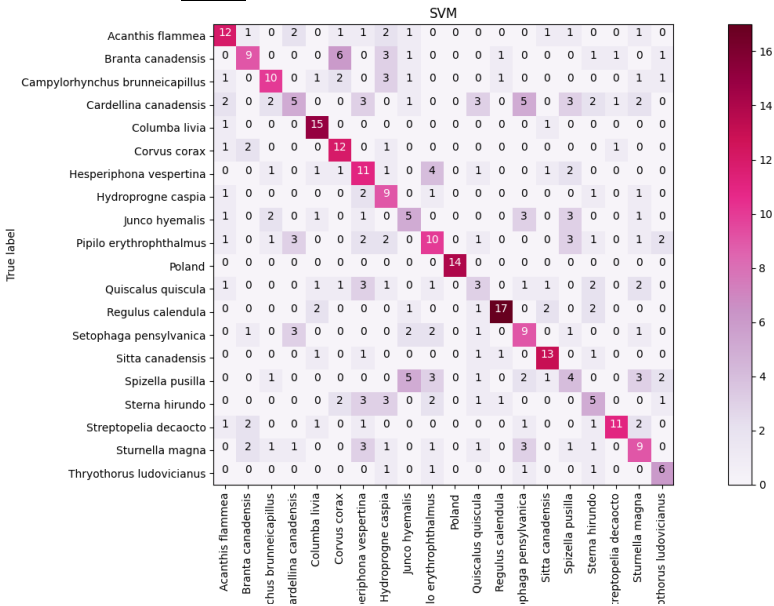
```
----- KNN - Results -----
Recall:  [0.34782609 0.43478261 0.28571429 0.34482759 0.58823529 0.58823529
0.30434783 0.33333333 0.23529412 0.33333333 1. 0.29411765
0.32 0.55 0.44444444 0.22727273 0.38888889 0.45
0.41666667 0.4 ]
Precision: [0.5 0.66666667 0.46153846 0.34482759 0.55555556 0.45454545
0.29166667 0.20833333 0.18181818 0.375 1. 0.22727273
0.72727273 0.35483871 0.38095238 0.33333333 0.31818182 0.47368421
0.41666667 0.28571429]
F1-Score: [0.41025641 0.52631579 0.35294118 0.34482759 0.57142857 0.51282051
0.29787234 0.25641026 0.20512821 0.35294118 1. 0.25641026
0.44444444 0.43137255 0.41025641 0.27027027 0.35 0.46153846
0.41666667 0.33333333]
Accuracy: 0.40 , 160
Number of samples: 400
```

ציפורים עם חזות זהה וקול דומה:

הציפורים Hydroprogne caspia וSterna hirundo מאוד דומות וראינו שהמסווג מתבלבל ביניהן לא מעט.



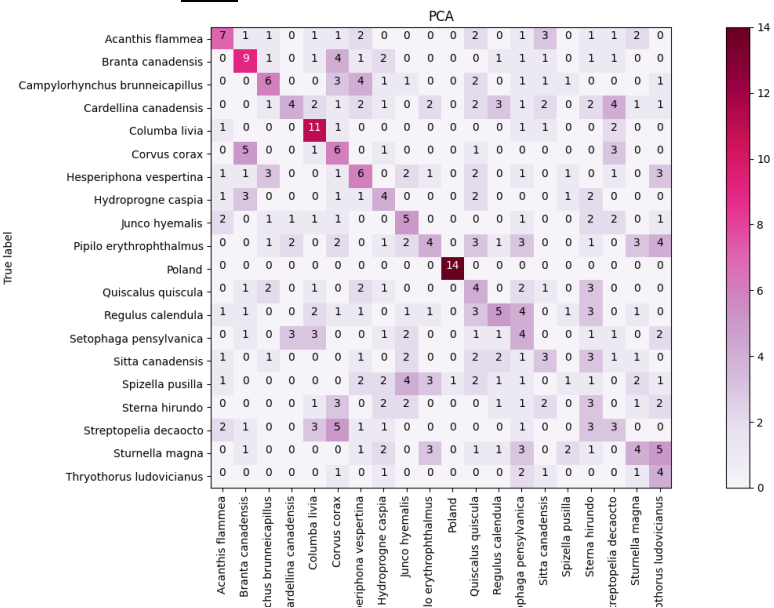
SVM:



```
----- SVM - Results -----
Recall:  [0.52 0.39 0.48 0.17 0.88 0.71 0.48 0.6 0.29 0.37 1. 0.18 0.68 0.45
0.72 0.18 0.28 0.55 0.38 0.6 ]
Precision: [0.55 0.53 0.56 0.36 0.65 0.48 0.35 0.33 0.29 0.4 1. 0.21 0.81 0.36
0.65 0.22 0.28 0.79 0.38 0.46]
F1-Score: [0.53 0.45 0.51 0.23 0.75 0.57 0.41 0.43 0.29 0.38 1. 0.19 0.74 0.4
0.68 0.2 0.28 0.65 0.38 0.52]
Accuracy: 0.47 , 189
Number of samples: 400
```



PCA:

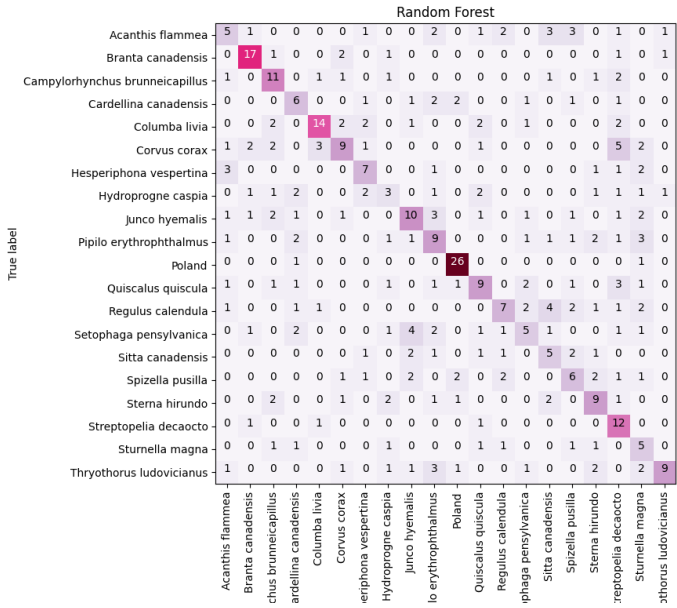


```
----- PCA - Results -----
Recall:  [0.3 0.39 0.29 0.14 0.65 0.35 0.26 0.27 0.29 0.15 1. 0.24 0.2 0.2
0.17 0.05 0.17 0.15 0.17 0.4 ]
Precision: [0.41 0.38 0.35 0.4 0.41 0.19 0.25 0.2 0.24 0.29 0.93 0.15 0.31 0.14
0.2 0.14 0.11 0.16 0.25 0.17]
F1-Score: [0.35 0.38 0.32 0.21 0.5 0.25 0.26 0.23 0.26 0.2 0.97 0.18 0.24 0.16
0.18 0.07 0.13 0.15 0.2 0.24]
Accuracy: 0.27 , 107
Number of samples: 400
```



• וקטור במימד 25 :

Random Forest:



----- Random Forest - Results -----

Recall: [0.25 0.74 0.58 0.4 0.54 0.35 0.47 0.19 0.4 0.39 0.93 0.41 0.32 0.25 0.36 0.33 0.47 0.8 0.42 0.41]

Precision: [0.33 0.71 0.48 0.35 0.7 0.5 0.44 0.25 0.45 0.35 0.79 0.45 0.5 0.36 0.29 0.33 0.43 0.34 0.22 0.75]

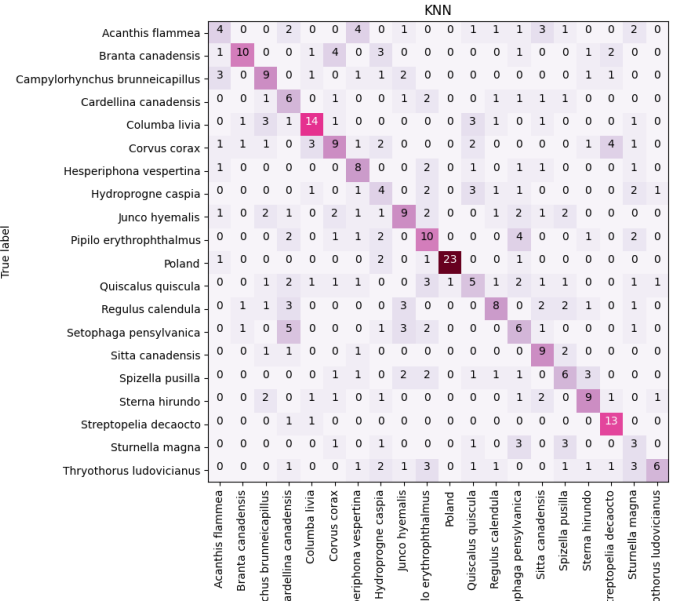
F1-Score: [0.29 0.72 0.52 0.38 0.61 0.41 0.45 0.21 0.43 0.37 0.85 0.43 0.39 0.29 0.32 0.33 0.45 0.48 0.29 0.53]

Accuracy: 0.46 , 184

Number of samples: 400



K-Neareast Neighbor:



----- KNN - Results -----

Recall: [0.2 0.43478261 0.47368421 0.4 0.53846154 0.34615385 0.53333333 0.25 0.36 0.43478261 0.82142857 0.22727273 0.36363636 0.3 0.64285714 0.33333333 0.47368421 0.86666667 0.25 0.27272727]

Precision: [0.33333333 0.71428571 0.42857143 0.24 0.60869565 0.40909091 0.38095238 0.2 0.40909091 0.34482759 0.95833333 0.27777778 0.5 0.24 0.40909091 0.31578947 0.5 0.59090909 0.16666667 0.66666667]

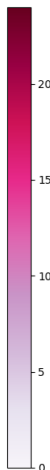
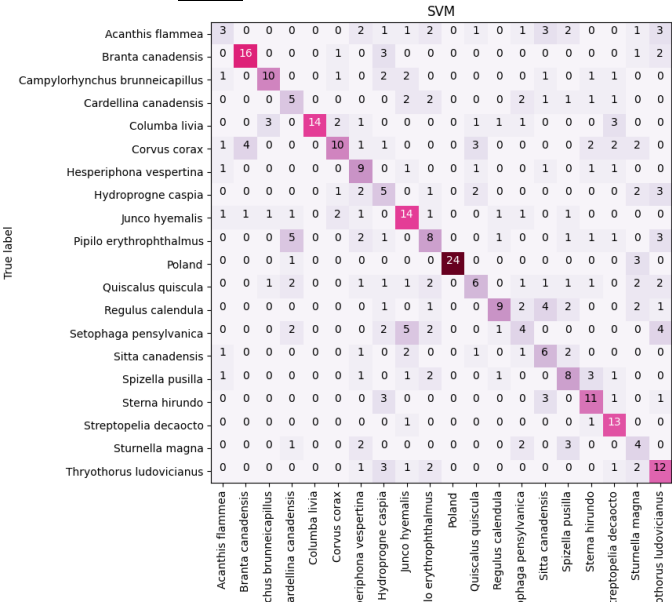
F1-Score: [0.25 0.54054054 0.45 0.3 0.57142857 0.375 0.44444444 0.22222222 0.38297872 0.38461538 0.88461538 0.25 0.42105263 0.26666667 0.5 0.32432432 0.48648649 0.7027027 0.2 0.38709677]

Accuracy: 0.43 , 171

Number of samples: 400



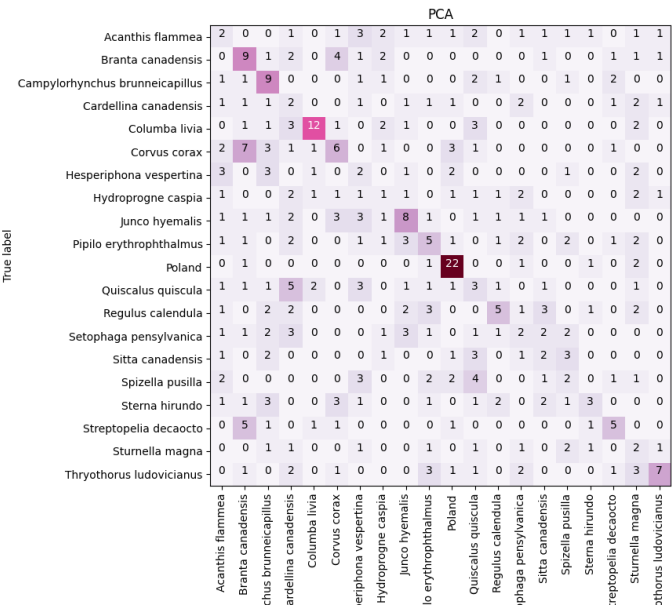
SVM:



```
----- SVM - Results -----
Recall:  [0.15 0.7  0.53 0.33 0.54 0.38 0.6  0.31 0.56 0.35 0.86 0.27 0.41 0.2
0.43 0.44 0.58 0.87 0.33 0.55]
Precision: [0.33 0.76 0.67 0.29 1.   0.59 0.38 0.22 0.45 0.35 1.   0.4  0.64 0.27
0.3  0.38 0.5  0.52 0.21 0.39]
F1-Score: [0.21 0.73 0.59 0.31 0.7  0.47 0.46 0.26 0.5  0.35 0.92 0.32 0.5  0.23
0.35 0.41 0.54 0.65 0.26 0.45]
Accuracy: 0.48 , 191
Number of samples: 400
```



PCA:

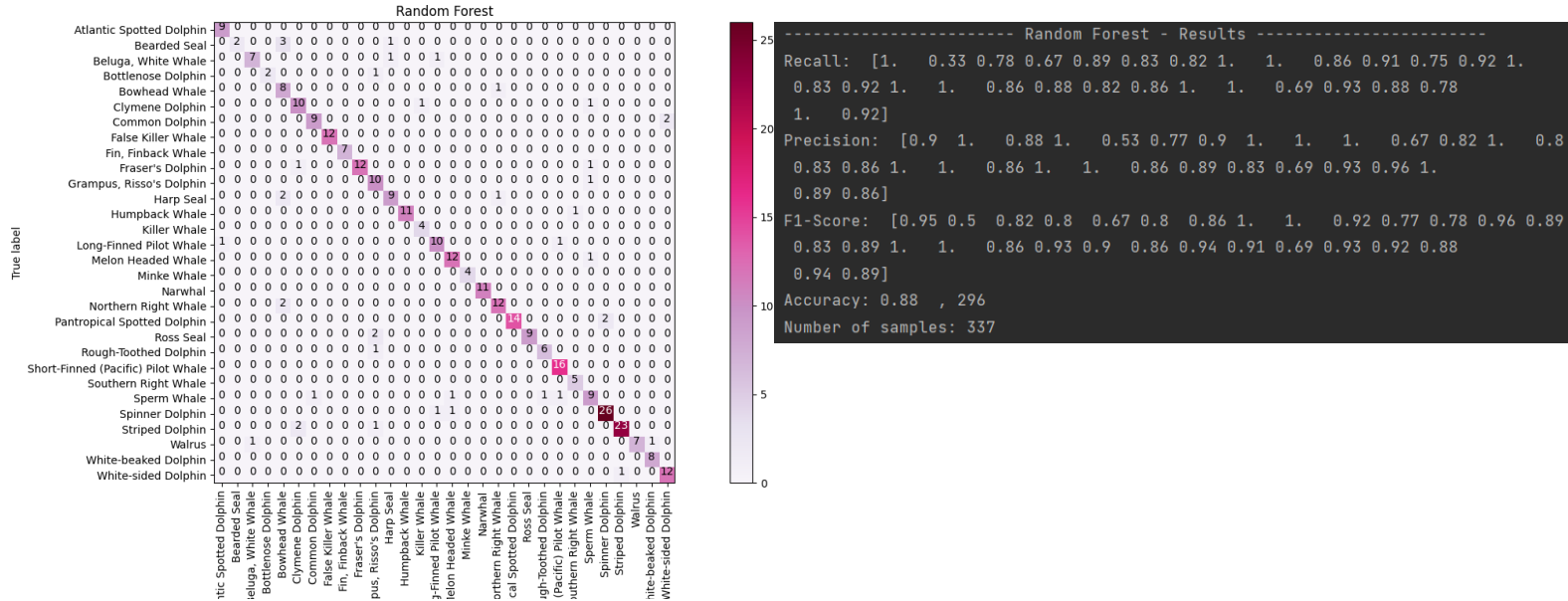


```
----- PCA - Results -----
Recall:  [0.1  0.39 0.47 0.13 0.46 0.23 0.13 0.06 0.32 0.22 0.79 0.14 0.23 0.1
0.14 0.11 0.16 0.33 0.17 0.32]
Precision: [0.11 0.29 0.29 0.07 0.67 0.29 0.1  0.08 0.36 0.24 0.59 0.12 0.38 0.12
0.14 0.13 0.38 0.38 0.09 0.58]
F1-Score: [0.1  0.33 0.36 0.09 0.55 0.26 0.11 0.07 0.34 0.23 0.68 0.13 0.29 0.11
0.14 0.12 0.22 0.36 0.11 0.41]
Accuracy: 0.27 , 109
Number of samples: 400
```

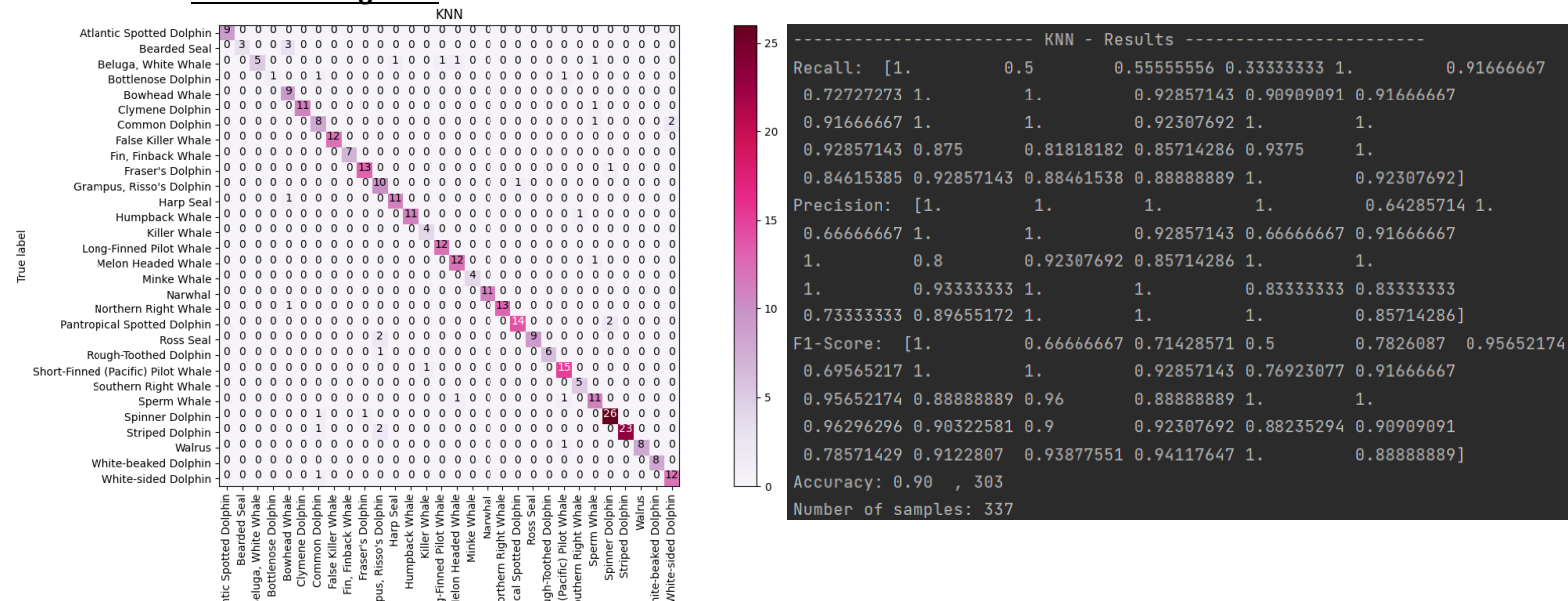
סיווג חיות ים:

• וקטור במימד 13 :

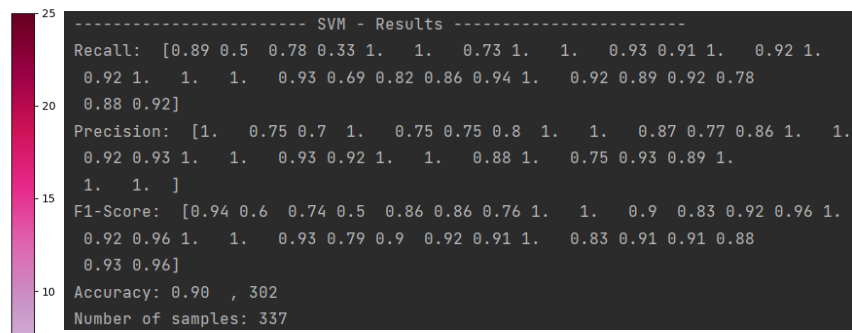
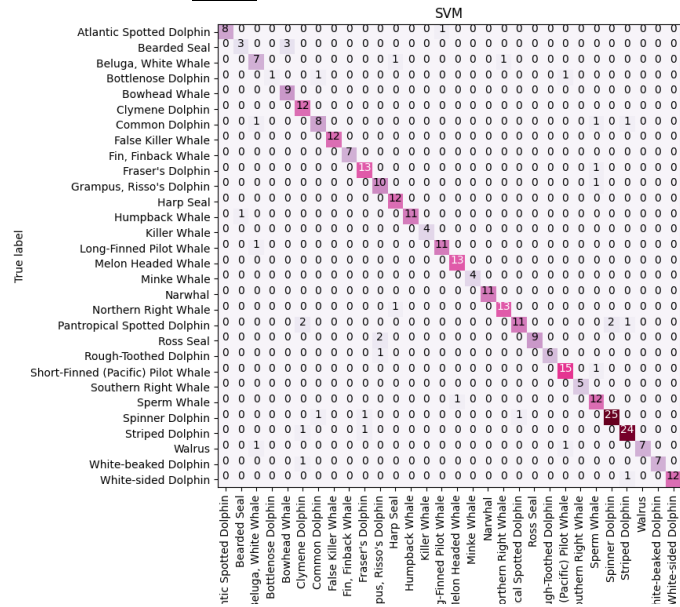
Random Forest:



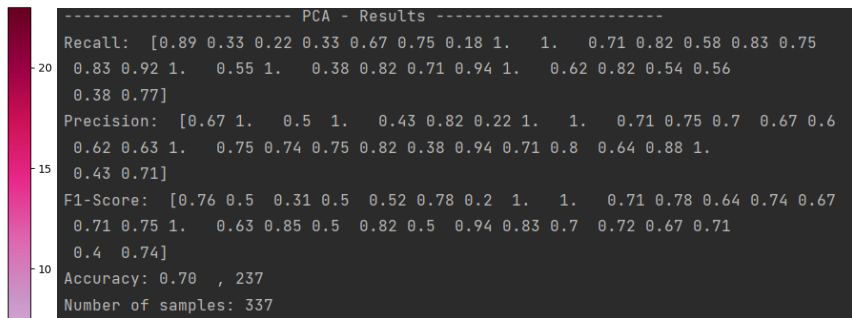
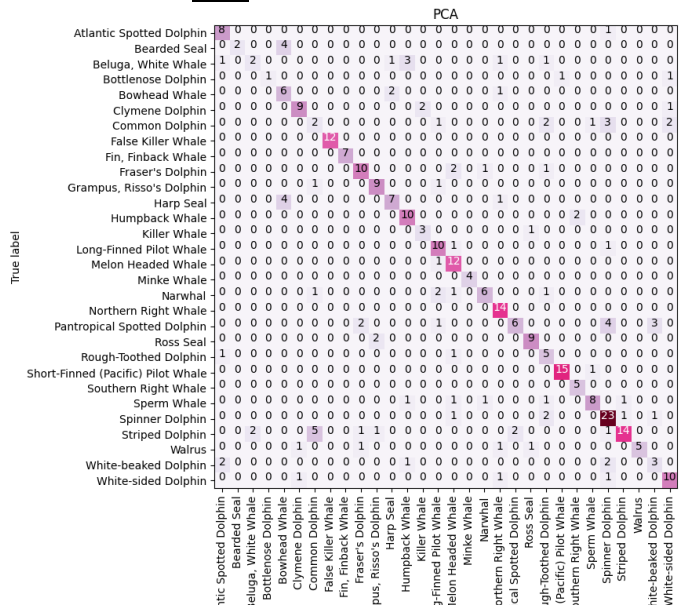
K-Nearest Neighbor:



SVM:

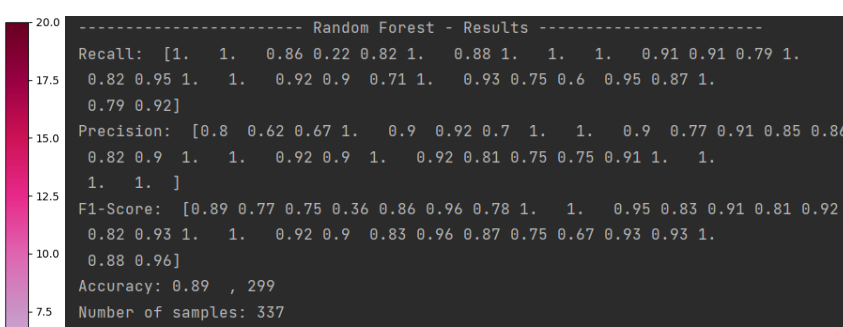
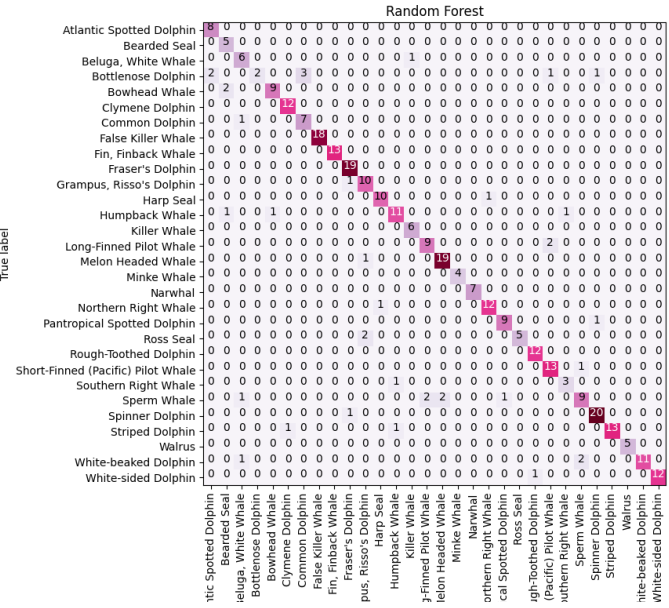


PCA:

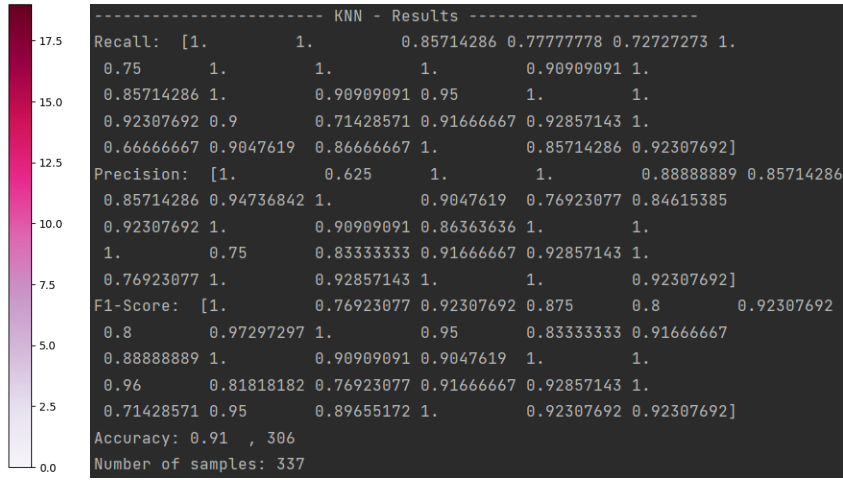
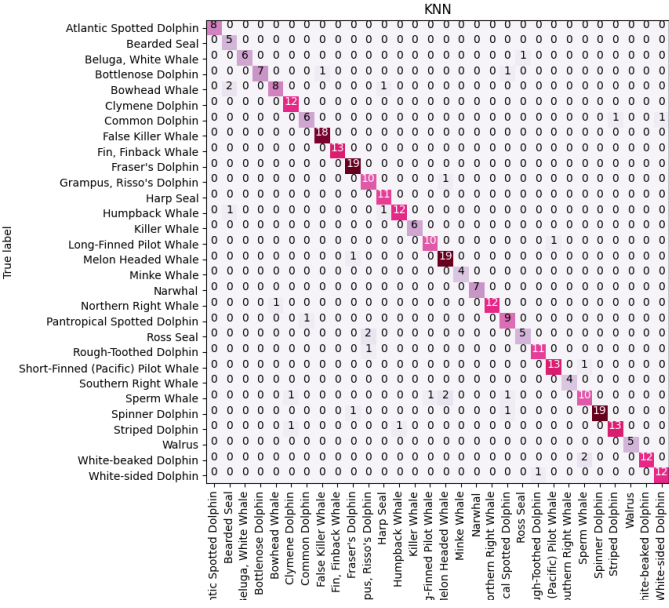


• וקטור במימד 25 :

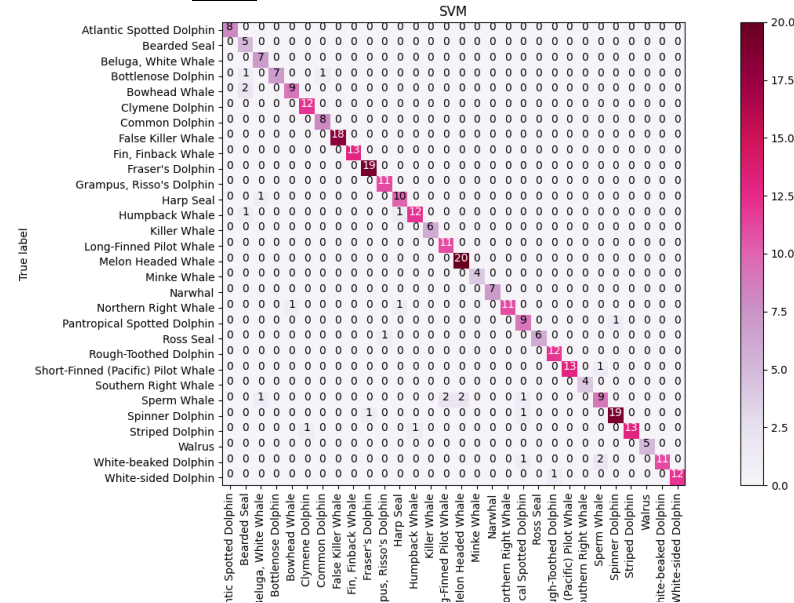
Random Forest:



K-Nearest Neighbor:



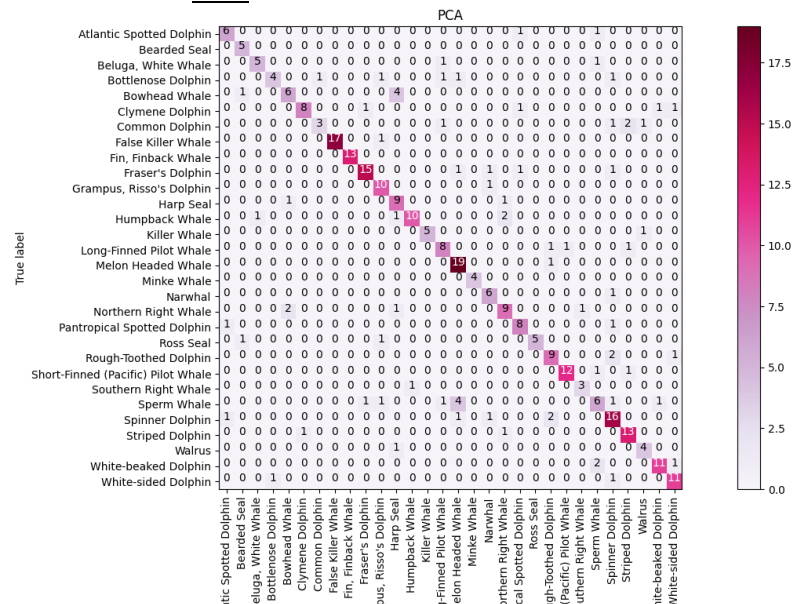
SVM:



```
----- SVM - Results -----
Recall: [1. 1. 1. 0.78 0.82 1. 1. 1. 1. 1. 1. 0.91 0.86 1.
1. 1. 1. 1. 0.85 0.9 0.86 1. 0.93 1. 0.6 0.9 0.87 1.
0.79 0.92]
Precision: [1. 0.56 0.78 1. 0.9 0.92 0.89 1. 1. 0.95 0.92 0.83 0.92 1.
0.85 0.91 1. 1. 1. 0.75 1. 0.92 1. 1. 0.75 0.95 1. 1.
1. 1. ]
F1-Score: [1. 0.71 0.88 0.88 0.86 0.96 0.94 1. 1. 0.97 0.96 0.87 0.89 1.
0.92 0.95 1. 1. 0.92 0.82 0.92 0.96 0.96 1. 0.67 0.93 0.93 1.
0.88 0.96]
Accuracy: 0.92 , 311
Number of samples: 337
```



PCA:



```
----- PCA - Results -----
Recall: [0.75 1. 0.71 0.44 0.55 0.67 0.38 0.94 1. 0.79 0.91 0.82 0.71 0.83
0.73 0.95 1. 0.86 0.69 0.8 0.71 0.75 0.86 0.75 0.4 0.76 0.87 0.8
0.79 0.85]
Precision: [0.75 0.71 0.83 0.8 0.67 0.89 0.75 1. 1. 0.88 0.71 0.56 0.91 1.
0.67 0.73 1. 0.67 0.69 0.73 1. 0.69 0.92 0.75 0.55 0.64 0.76 0.67
0.85 0.79]
F1-Score: [0.75 0.83 0.77 0.57 0.6 0.76 0.5 0.97 1. 0.83 0.8 0.67 0.8 0.91
0.7 0.83 1. 0.75 0.69 0.76 0.83 0.72 0.89 0.75 0.46 0.7 0.81 0.73
0.81 0.81]
Accuracy: 0.77 , 260
Number of samples: 337
```




סיכום תוצאות – טכניקות שעבדו יותר טוב או פחות טוב :

בכל הסיווגים האלגוריתמים Random Forest, K-Nearest Neighbor, SVM, עבדו טוב ונתנו תוצאות דומות, שנעו בטווח קטן, לעומת זאת, בשימוש בPCA ניסינו להקטין את המימד ל-2 וקיבלנו תוצאות מאוד לא טובות, לכן הגדלנו אותו ל-5, התוצאות שקיבלנו עדיין לא היו טובות אבל ביחס להקטנה המשמעותית (יותר מחצי) הן סבירות, גם כשהשתמשנו בווקטור ממימד 13 וגם כשהשתמשנו בווקטור ממימד 25 הקטנו את המימד בPCA ל-5 וקיבלנו תוצאות טובות יותר עם הווקטור ממימד 25 למרות ששינוי המימד הוא משמעותי וגדול יותר.

בנוסף, לרוב קיבלנו תוצאות קצת יותר טובות בשימוש עם וקטור ממימד 25 ולעיתים התוצאות היו זהות.

סיווג חיות ים – משולב:

בסיווג החיות עם הקלטות משולבות לא קיבלנו תוצאות טובות, האלגוריתמים הצליחו לסווג חיה אחת מתוך 2 שמיזגנו להקלטה אחת ולעיתים גם לא הצליחו לזהות כלל, גם כאשר ניסינו לקחת חיות מאוד שונות התוצאות לא היו מספקות, לכן הסקנו שמיזוג ההקלטות באופן מלאכותי לא מאפשר סיווג טוב ושיערנו שהקלטה משולבת מקורית תניב תוצאות טובות יותר.

	Fraser's Dolphin combined with Grampus, Risso's Dolphin	Harp Seal combined with Beluga, White Whale	Spinner Dolphin combined with Striped Dolphin
Random Forest	סיווג חיה אחת נכון - Grampus, Risso's Dolphin	סיווג חיה אחת נכון - Beluga, White Whale	לא סיווג נכון אף חיה
SVM	סיווג חיה אחת נכון - Grampus, Risso's Dolphin	סיווג חיה אחת נכון - Beluga, White Whale	לא סיווג נכון אף חיה
K-Nearest Neighbor	סיווג חיה אחת נכון - Fraser's Dolphin	סיווג חיה אחת נכון - Beluga, White Whale	לא סיווג נכון אף חיה
Image			

אתגרים:

1. סיווג מאגר הציפורים:

סיווג מאגר הציפורים היה מאתגר יותר מסיווג חיות הים, המאגר שמצאנו לא היה מאוזן ולכן לא הניב תוצאות טובות בכלל (קיבלנו דיוק של 10-20 אחוזים), נאלצנו לאזן את המאגר ולהוריד ציפורים עם מספר הקלטות קטן על מנת להעלות את רמת הדיוק, המאגר כרגע עומד על בערך 40-50 אחוזי דיוק, ניסינו להבין ממה זה נובע, ובדקנו את זני הציפורים במאגר, מצאנו 2 זוגות של ציפורים בעלות המון תכונות חיצוניות זהות וגילינו שיש המון טעויות בסיווג עליהן. הסקנו מכך שיש לנו במאגר ציפורים עם זנים מאוד דומים ומכאן הטעויות בסיווג. בנוסף קראנו קצת על המאגרים והבנו שלכל זן יש סוגי הקלטות שונות (שירה, קריאת אזהרה, קריאה לעזרה, ועוד). אנחנו חושבות שעצם ההבדל בין ההקלטות בכל זן מקשה על הסיווג.

2. טעינת הקבצים:

המאגר מכיל כמעט 3700 קטעי קול באורכים שונים, לכן טעינת הקבצים לקחה לנו כמעט שעתיים של זמן ריצה, הבנו שלא נוכל לעבוד כך בכל פעם שנריץ את אלגוריתמי הסיווג ולכן טענו את הקבצים בפעם האחת וייצאנו את הווקטורים לקובץ טקסט. כך יכולנו לקרוא את הווקטורים מקובץ טקסט ולהתגבר על בעיית הזמן.

3. הקלטות משולבות:

רצינו לבדוק אם נוכל לסווג קטעי קול בהם נשמעים 2 זנים של חיות בו זמנית, מתוך הבנה שהן לא חיות בנפרד ונרצה להבדיל בניהן גם כשהן משמיעות קול בו זמנית. לא מצאנו מאגר עם קולות משולבים לכן יצרנו בעצמנו מספר הקלטות באמצעות תוכנה לעריכת קטעי קול (audacity), חיפשנו במאגרים הקלטות בעלות אותו אורך בהן שומעים את קול החיות בבירור ומיזגנו אותם להקלטה אחת שאותה ניסינו לסווג.