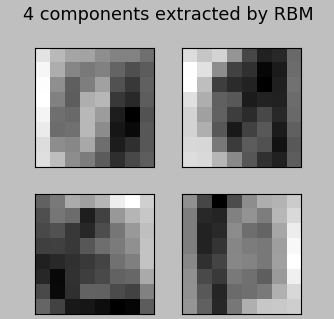
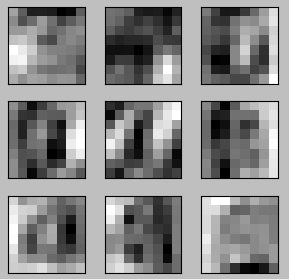
**תרגיל 3 – מערכות לומדות – מגיש: נתאי יחזקאלי, 021976808**

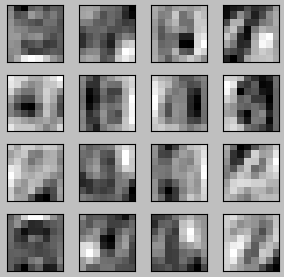
חלק ראשון  
שאלה 1:

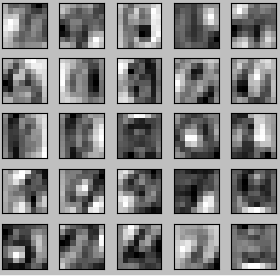
ב.

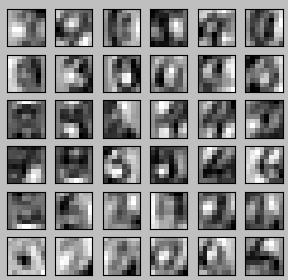
תצוגות גרפיות:

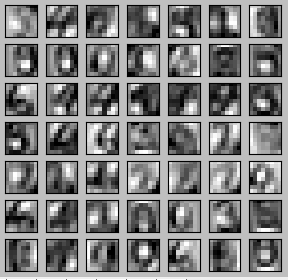










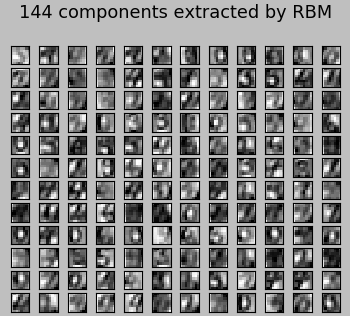


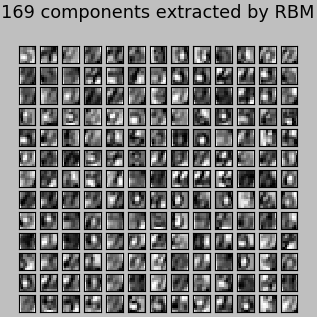


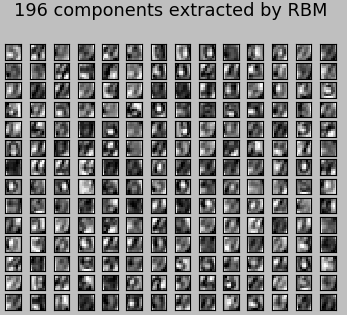


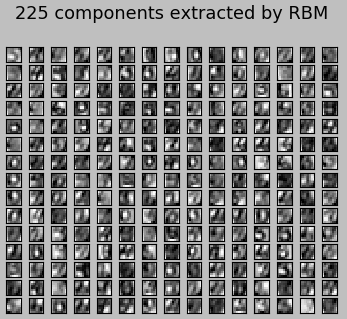


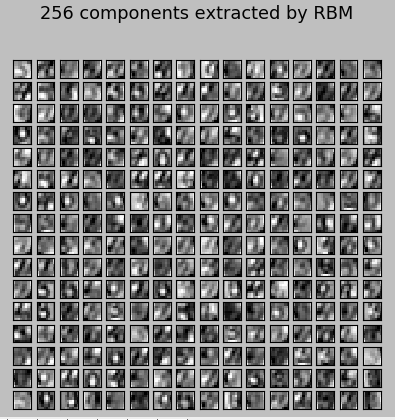


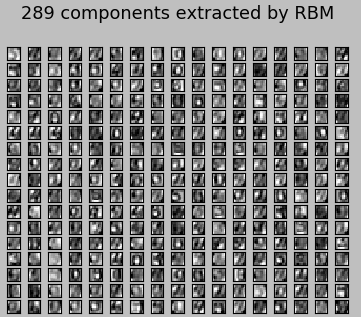
התחלתי להוסיף גם את הכותרת כי אחרת ממש קשה לספור את הכמות:

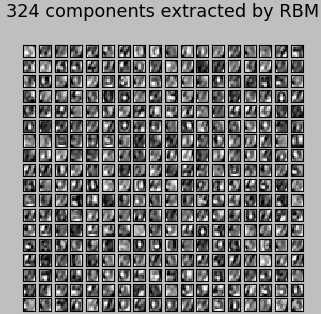


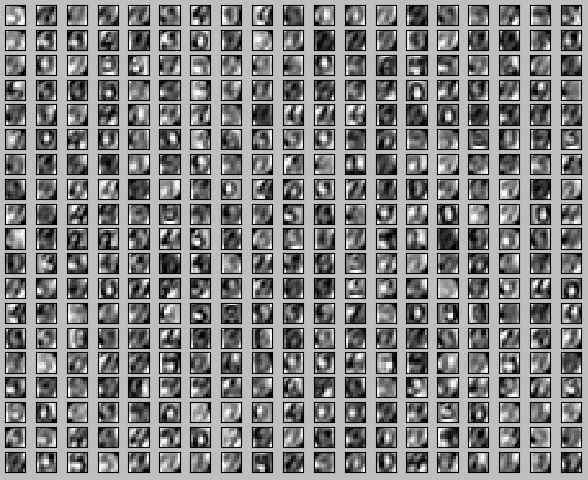


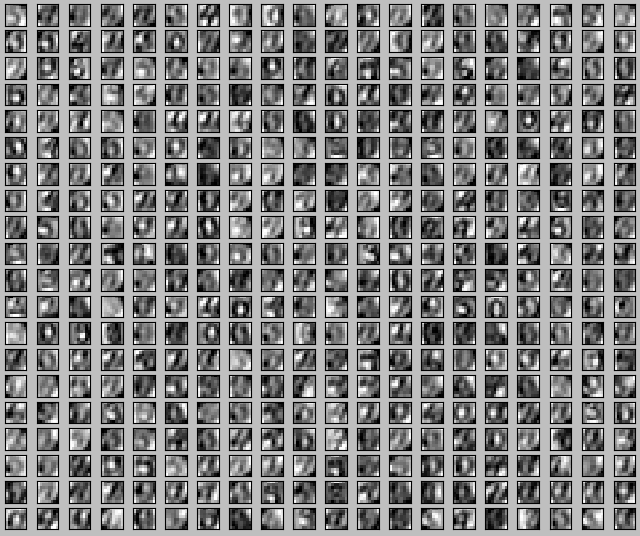






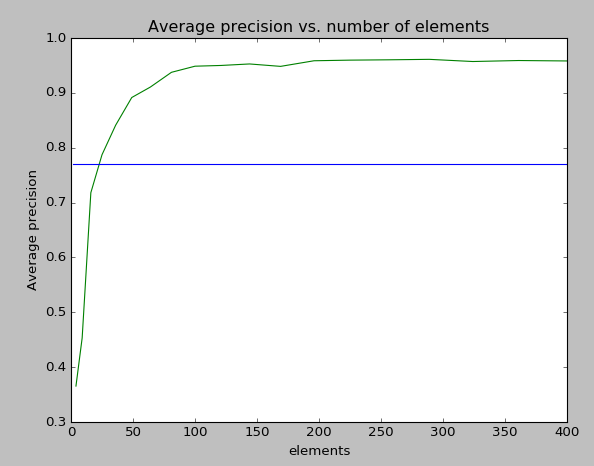


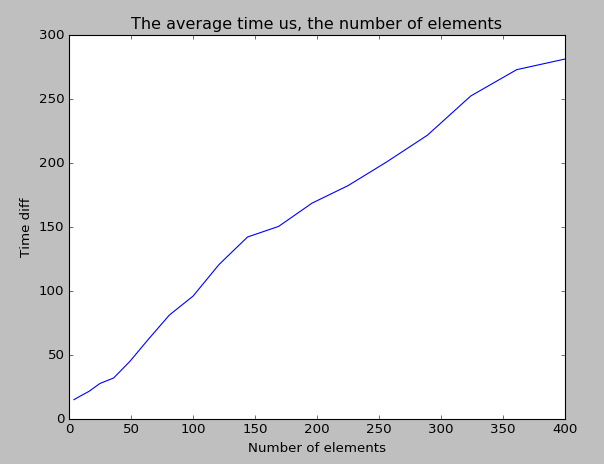
בגלל המרווח בין הכותרת לתמונות החלטתי לוותר מעכשיו (זה סה"כ עוד שתי תמונות) על הכותרת כדי שהתמונות יכנסו לדף:



ג.

והנה שני הגרפים (ניתן לדעת מה מודד כל גרף לפי הכותרות והצירים): *הערה: בגרפים כתוב elements אבל הכוונה היא ל-****components*** *(גיליתי את זה רק אחרי שסיימתי את כל התרגיל, לכן אמרתי לעצמי שאין טעם לשנות (כי זה יכול לקחת המון להריץ את כל) אלא כדאי לכתוב את זה בהערה).*





ד.

1.

המימדים של המשתנים הם:

|  |  |  |
| --- | --- | --- |
| X\_train | בכל הרצה המימד הוא: | כלומר, X\_train מכיל 7188 מערכים שכל אחד בגודל 64. |
| X\_test | בכל הרצה המימד הוא: | כלומר, X\_test מכיל 1797 מערכים שכל אחד בגודל 64. |
| rbm.transform(X\_train) | בהרצה המימד הוא: | כלומר, בהרצה הוא מכיל 7188 מערכים מגודל |
| rbm.intercept\_hidden\_ | בהרצה המימד הוא: | כלומר, בהרצה הוא מכיל מערך בגודל |

הסבר על הגודל/מימד (שנמצא בטבלה למעלה): – פירושו שאותו משתנה מכיל מערכים שגודל כל אחד מהם הוא , במידה ואין אזי זה אומר שיש לנו מערך אחד בגודל .  
ולכן אנחנו יכולים לכפול משהו כגון: .

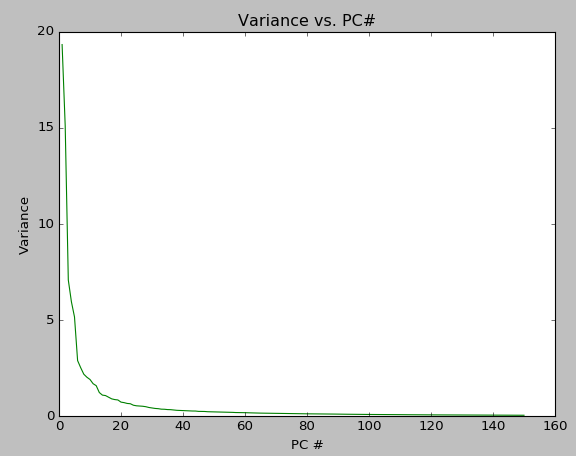
2.

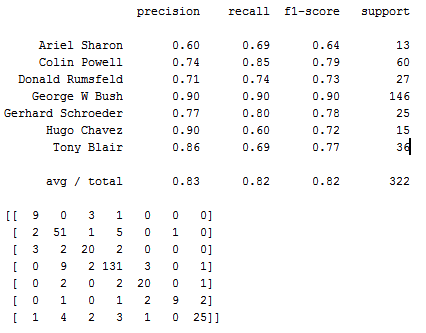
אני פשוט אכתוב את הקוד עם המשתנים המתאימים ולמטה את החישוב הסופי.

**def** transform(self, X\_train):  
 *"""Compute the hidden layer activation probabilities, P(h=1|v=X).  
  
 Parameters  
 ----------  
 X : {array-like, sparse matrix} shape (n\_samples, n\_features)  
 The data to be transformed.  
  
 Returns  
 -------  
 h : array, shape (n\_samples, n\_components)* ***n\_components = i*** *Latent representations of the data.  
 """* check\_is\_fitted(self, **"components\_"**)  
  
 X = check\_array(X, accept\_sparse=**'csr'**, dtype=np.float64)  
 **return** self.\_mean\_hiddens(X\_train)  
  
**def** \_mean\_hiddens(self, X\_train):  
 *"""Computes the probabilities P(h=1|v).  
  
 Parameters  
 ----------  
 v : array-like, shape (n\_samples, n\_features)  
 Values of the visible layer.  
  
 Returns  
 -------  
 h : array-like, shape (n\_samples, n\_components)  
 Corresponding mean field values for the hidden layer.  
 """* p = safe\_sparse\_dot(X\_train, (64,i\*\*2))🡨🡨 כאן מתבצע הכפל  
 p += (i\*\*2,)  
 **return** expit(p, out=p)

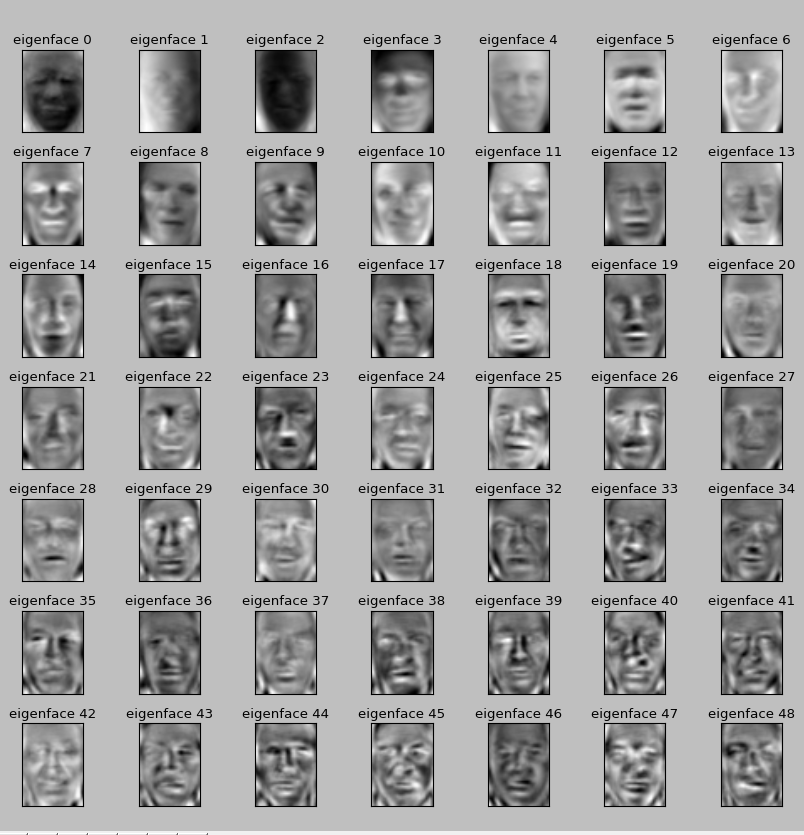
המימדים של המשתנים אכן מתאימים לחישוב כי אם נסתכל על הגודל של המימדים נראה שנקבל:p = safe\_sparse\_dot((7188,64), (64,i\*\*2)) ואכן נוכל לכפול שני מערכים בגודל כזה. מה שנקבל יהיה בגודל .

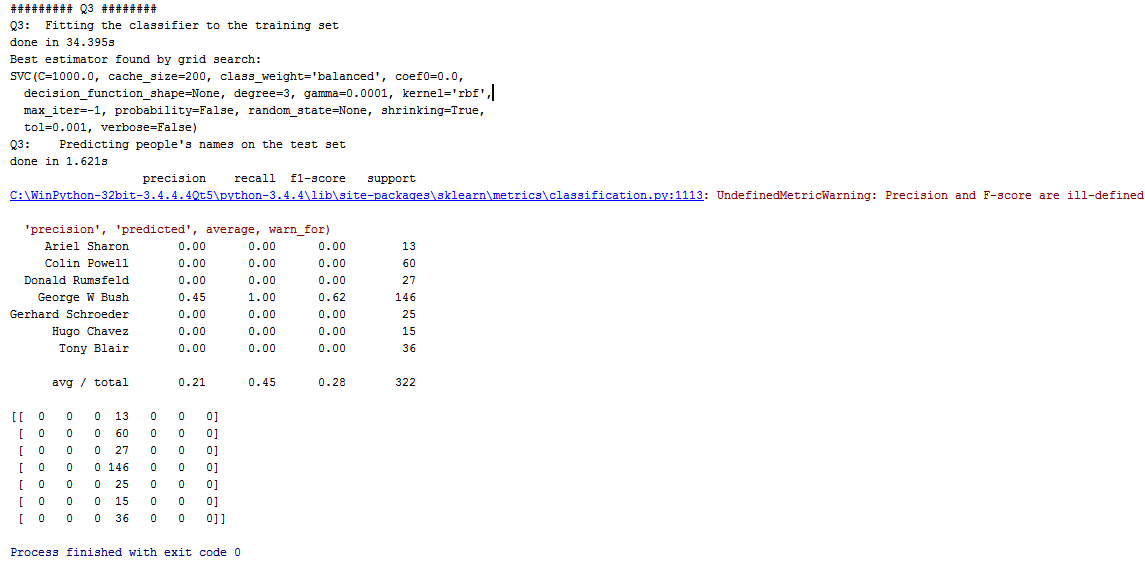
חלק שני

1. 

לגבי הסעיף הזה לא ממש הבנתי מה אני אמור לשים בו, כי הזמן והתוצאה נתונים לי בפלט:  
זמן: done in 39.159s.  
התוצאה:  


1. (לא הבנתי אם צריך להגיש את החלק הזה או לא אז הוספתי אותו לכל מקרה)



ה. 

הדבקתי את הפלט כתמונה כי אחרת זה יוצא בצורה בעייתית....

(החצים מורים על הזמן).

ו. אכן. הפחתת המימדים אכן הועילה (הפחתת המימדים מ-1850 ל-150 שזה פחות מעשירית). ניתן לראות זה כי בסעיף הקודם כל התוצאות סווגו כאותו הדבר. הסיבה לכך היא שכשאר יש כמות מימדים גדולה צריך כמות עצומה של נתונים בשביל יצור ניבוי מוצלח, והיות ואין לנו כמות גדולה של נתונים (יש לנו יותר מימדים מנתונים, כאשר אין שום הפחתה, מה שמקשה מאוד על המערכת ליצור ניבוי נכון). לכן יצא לנו שהמערכת תייגה את כל הנתונים כנתון שמופיע הכי הרבה פעמים במאגר (כמעט 50%), כי המערכת למדה אותו הכי טוב ולכן התאימה את הכל אליו.  
לגבי השאלה הראשונה (בתרגיל בכלל), שם ניתן לראות שהפחתת המימדים לא עזרה לנו (לפי מה שכתוב בדף התרגיל – מספר הרכיבים=מספר המימדים), כי במקרה הנ"ל – מספר המימדים היה קטן ביחד למספר הנתונים (מספר המימדים המקורי היה 64 ומספר הנתונים היה גדול מ-7000 [אני מתייחס לסט האימון]) כלומר, מספר הנתונים היה בערך פי מאה גדול יותר מאשר מספר המימדים, ולכן כאן – הוספת הנתונים/הפחתת מימדים לא באמת עזרה לנו וניתן לראות את זה גם בתוצאות.

**הערות לתרגיל:**

יורם, יש לי כמה הערות שאשמח מאוד אם תתיחס אליהן:

1. בפונקציות לא שמתי תיעוד ברובן. היות והן התשובות לשאלות בתרגיל.
2. היות והיה לי בילבול עם השאלות – אזי כאן בדף נתתי עניתי על הסעיפים לפי מספרם, אבל בקבצי קוד פשוט מספרתי את התשובות החל מ-1. אחרת זה היה מבלבל אותי. אודה לך אם תוכל להתחשב בכך...
3. בשאלה השנייה (בחלק השני), יש איזו אזהרה. בגלל חוסר הזמן לא יצא להתעמק ולתקן אותה. אודה לך אם תתעלם ממנה.
4. שים לב שבגרפים של שאלה 1 יש בילבול עם אחת המילים (כתבתי גם הערה על כך). לוקח המון זמן להריץ את התוכנה עד שהיא מסיימת, לכן במקום להריץ שוב החלטתי לכתוב את ההערה (בקוד עצמו המילה תוקנה).
5. בחלק השני – בחלק של ה-PCA לא שמתי את הערכים הטובים של ה-Grid Search היות ורציתי לראות כמה זמן לוקח לזה לבדוק מה הכי טוב. לעומת זאת – בחלק השני – ללא ה-PCA – זה לקח המון המון זמן עד שזה מצא את הערכים הכי טובים, כך שהחלטתי שבמקרה הזה אני כבר כן אצמצם את המערכים לערך אחד (אם זאת לא רשימה הוא מחזיר שגיאה).

תודה רבה,

נתאי