

אסכולה נמוכה (קלאסרי צ'י'קן) זה מהד'ק  
 סיווג אובייקטים לקלאסים של אובייקטים  
 צומ'ם.

מסמכים מאומדן אסכולה חייבים להיות  
 צומ'ם ומסמכים מאסכולה שונים חייבים  
 להיות שונים.

קלאסליצ'יה זה לא מהד'ק סיווג כמו שאמחנו רח'ם  
 כ' כאן אין לנו להשתמש בנתונים מסומנים  
 לקבוצות מיוחדים וחוצה למחצה של מהד'ם ק"ח'ם.  
 יש לנו כמה סוגים של אלגוריתמים של

קלאסליצ'יה. כמו Hierarchical, Partitional, Flat.

לראשון אלו יכולים להיות K-means אלגוריתם.  
 זה אלגוריתם אויוריטי שבו אמחנו בוחרים  
 כמה אסכוליות להשתמש; כלומר כמה קבוצות  
 ע'צור, א זה מס' הקבוצות. אם א סגור יותר  
 אם כמה האל'ים הקבוצה קטן. פל' של אלגוריתם  
 זה הוא קבוצת labels.

ה-מט' אלגוריתם אמחנו ממח'ים עם כמה מסויים  
 של אסכוליות, הוא גם תת'אלגוריתם של K-means.

בש'ל זה אלו מק'לים כמה האסכוליות ע"י  
 מ'צ'ם קבוצות ע' שני'ע' לקבוצה אח' סגורה,  
 שמכילה כל האובייקטים בה. ש'ל' פל'ת ממח'ת  
 במשנה מרחק בין האובייקטים שבה'ל זה יש  
 לנו כמה אלגוריתמים שונים כמו קוס'ינוס,  
 אוקלידי ומנה'ן.



1) בספרייה sklearn יש מודל של אלמנטים

k-means שקרה means. אלמנטים זה צריך לקבל מס' של אסכוליות שאיתו 'צריך' לעבוד ואילו להימנע וצריך להקל'ן עדיק הנקרה inertia המ"צם ע"כ כמה אסכוליות יכולים להיות שווים ודב'נים. במקרים רבים יש לנו מודל הנקרא mini batch k-means ש"ה לטף עובדת עם נתונים הממוצעים לכמה קבוצות. כ"ו להוסיף עומס ולחשב נתונים יותר מהר. זה משאז ש"חוש' אלהם לפתור פוגע בתוצאות. אנו יכולים להצביע אלה התוצאות האלו עם שני צ'רים הו"ה"ו אולם תקוצות נפופות, והמחלוקה נתונים שלנו.

קל"אס' ה"נ"ם הי"ארכי קטן אלה האלמנטים ה"ח'ים המ"ים ש"פ"ה מ"יו"ם אסכוליות אלו ע"פ, מ"יו"ם ה"י"א, ה"ה מ"מ"מ"ם, ה"י"א זה מ"מ"מ"ם כ"ר"ל ש"ל צ"נ"צ"מ"ה זה ה"ש"ר הע"יו"ן ה"ו"ל ה"ש"כ"ל ה"כ"ל"י ש"כ"ל ה"ו כל הנתונים. אלמנטים זה מ"מ"מ"ם

ה"י"א bottom up זה מ"מ"מ"ם מ"מ"מ"ם מ"ה ש"כ"ל א"ב"י"ק ק"י"ם א"ש"כ"ל מ"ש"ל ו"ע"ו"ם מ"יו"ם ע"ש"ר ה"א"ש"כ"ל"י"ם ו"כ"ה מ"י"ע"ים ע"ש"כ"ל א"ח"ז ו"י"ח"ז ש"מ"כ"ל מ"ע"ל כ"ר"ל ע"מ"ד"ל זה יש מודל מקביל הנקרה

Feature Agglomeration. ה"ט מ"מ"מ"ם ה"א"ש"כ"ל"י"ם

צ"מ"ו"ר כ"י מ"מ"ר א"ח"מ"ם כ"י"ח"ז א"ש"כ"ל א"ח"ז.

ע"ק"ח"ם כ"א"ל"ו"ם צ"מ"ם ו"מ"ח"ב"ים כ"ק ש"כ"ל יש ל"נו כ"מ"ו ק"ל"ה ו"ח"ר ו"ל"ו מ"ד"צ' ל"נו א"ל- ו"כ"מ"ו ה"מ"מ"ד"ים,



2) הסבר ה K-MEANS  
למשל:  $k$ -means clustering. כולל בו  $k$  clusters,  $k$ -means,  $k$ -group average agglomerative.  
כמו בספר"ה קוצות מחיל מוזכר  
מס' א באלגוריתם ואז בוחר כל וקטור עם  
ערך הכ קרוב לאלגוריתם ואז מחשב  
מרכז ה, אלגוריתם סקורה centroid. מבצעים התהליך  
מחירי גמול ושינוי שלהם לאלגוריתם עד שמת  
אלו כולם ויהיו לנו אלגוריתם סופי.

אלגוריתם GAAL מחיל משיוך כל אלגוריתם  
לאלגוריתם אחר משל כך שמתחיל האלגוריתם יהיה שווה  
למחירי הוקטורים. ואז בשיטה איטרטיבית נעשה  
מיצוי של כל האלגוריתם הכי קרובות עד שלבסוף  
לאלגוריתם אחר כללי עשוי. התהליך הזה נמשך לנו  
הסוף עד של נצטרף בו נכנס למחירי אחר  
כל התהליך המינימום של אלגוריתם מאלגוריתם של וקטורים  
עד השורה של העד.

אלגוריתם Gaussian EM גורם וקטורים ע"י  
ה Gaussian א מקוריות. ומקוריות האלה הם  
הסתברות של אופיי, הממוצע ומאריצט. ומשמים  
הם כז' למצוא הסתברות של גמול  
סקיבלו. קוצת מחסלים והסתברות של כל  
וקטור באלגוריתם ואז פתאום של פסוק ממוצעים  
ע"י הסתברות מקסימלית שנקחה מהסתברות  
של אלגוריתם ומתחילת כל מחסל עד שהסתברות של  
גמול לא מחסל עצומה.



## סיכום

המעבדה זו הושתנה להכיר את הספריות של  
פ"אן כמו אדלמ ו- Harsanyi והכנו יותר  
מונחים ומכונות של אלגוריתמי אשכול וצבירה  
עם עיבוד לקס.

השתמשו בעיבוד מקדים של קבצי לקס  
כמו שלמדנו במעבדות קודמות ובחרנו לעבוד  
עם לקס מומחה. השתמשו באלגוריתם  $k$ -means  
לגילוי קלאוסטרים וכן עמדנו עם המונחים בהמשך.  
עשינו מספר ניסויים עם המונחים ואלגוריתמים  
לאשכול גמורים. כל פעם ביקרנו במובאות עבור  
כמות אשכולות שונים ובחרנו בצורה יותר טובה  
של נבואה. שינינו כל פעם מאלגוריתם המחקר  
בין התקדמות מידע לבדוק ניסוי.

למדנו לנהל סדרים של אשכול וצבירות  
גמורים לפי צורך וספריות השימושים.  
עשינו השוואה בין לחי לחי של מערכות וצורה  
המונחים עם גמור בדיקה ואיתון. והגענו  
למסקנה של אלגוריתמי Harsanyi יותר יעילים  
בנייה מודל אשכול וסם במובאות בסכומים  
Harsanyi למרות ש- אדלמ אלגוריתם  
היארי עבדו יותר מהר אבל מובאות היו לסם  
ממש גומות.

אחרי כל הניסויים הוסנו למסקנה שבהנייה  
מודל צריך להיות שזו כמה חסמה לנו  
הנבואה של לחי לחי ולא עמדו לפי כל אלה  
בכל מקרה Harsanyi הביא מובאות יותר יעילות.