**חישוב של מבחנים סטטסיטים של כמה אלגוריתמים על כמה data-set:**

**יש 2 שלבים למבחנים הסטטיסטים:**

1. **נבדוק אם הביצועים של כל האלגוריתמים על כל ה-data\_set זהים או לא.**
2. השערת H0: על האלגוריתמים זהים

השערת H1: לא כל האלגוריתמים זהים.

1. נבדוק את המדד שמעניין אותנו עבור כל אלגוריתם עבור כל data-set, המדד הקלאסי זה accuracy, אבל אנחנו יכולים להשתמש בכל מדד של clusters שמעניין אותנו. נכניס לטבלה:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  | **Alg3** | **Alg2** | **Alg1** |  |
|  |  |  |  |  |  | **96.3** | **98.3** | **Data\_set1** |
|  |  |  |  |  |  | **93.2** | **92.1** | **Data\_set2** |
|  |  |  |  |  |  |  |  | **......** |

1. ניצור טבלה שמתסכלת על המיקומים של כל אלגוריתם בכל data\_set, לא נסתכל על תוצאות הדיוק בפועל, אלא רק על המקום בדירוג:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  | **Alg3** | **Alg2** | **Alg1** |  |
|  |  |  |  |  |  | **2** | **1** | **Data\_set1** |
|  |  |  |  |  |  | **1** | **2** | **Data\_set2** |
|  |  |  |  |  |  |  |  | **......** |
|  |  |  |  |  |  |  |  |  |

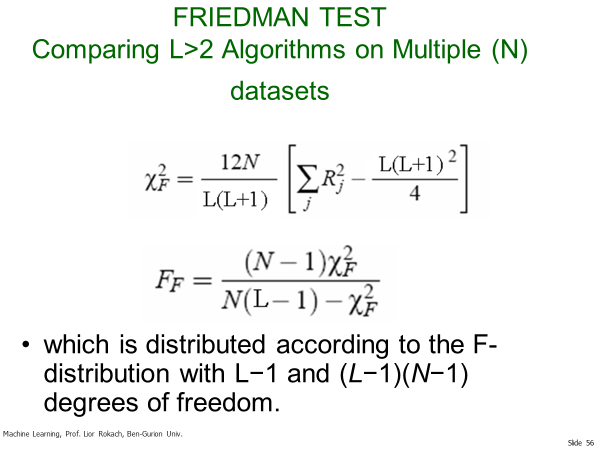
1. נסכום עבור כל אלגוריתם את סכום הדירוגים

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  | **Alg3** | **Alg2** | **Alg1** |  |
|  |  |  |  |  |  | **2** | **1** | **Data\_set1** |
|  |  |  |  |  |  | **1** | **2** | **Data\_set2** |
|  |  |  |  |  |  |  |  | **......** |
|  |  |  |  |  |  |  |  |  |

1. נשתמש בנוסחא המצורפת בשקף הבא בשביל לחשב את הסטטיסטי:

L- מספר האלגוריתמים

N- מספר ה-data set.

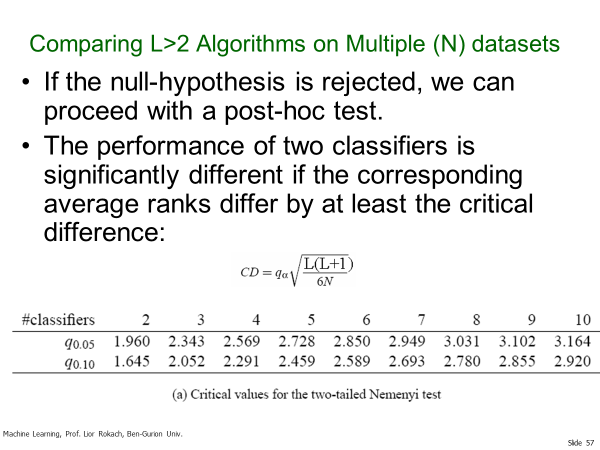
****

1. נשווה את הסטטיסטי שהתקבל בסעיף ד' למול הקריטי שאותו נוציא מטבלת חי בריבוע, עם L-1 דרגות חופש. סטטיסטי גדול מקריטי נדחה את השערת ה-0, אחרת לא נדחה.

2 . במידה ודחינו את השערת ה-0 בחלק הקודם, קיימים 2 מקרים אפשריים:

1. יש כמה אלגוריתמים ונרצה להשוואות בין כל זוג, מי יותר טוב והאם ההבדל בניהם מובהק סטטיסטית.
2. יש אלגוריתם חדש שפיתחנו שנרצה להשוואות את כולם מולו.

**מקרה א:**



עבור כל זוג אלגוריתמים שנרצה להשוואות נגיד 1 ו-4, נראה אם האפשר בניהם גדול מערך הסף CD, אם ההפרש בין שניהם יותר גדול מ-CD, אפשר להגיד שההפרש בניהם מובהק סטטיסטית.

אם ההפרש בניהם קטן מ-0, אז זה אומר ש- R1 מובהק סטטיסטית יותר טוב מ- R4 (ככל שהמספר קטן יותר זה יותר טוב.

מקרה ב':

