

Applied Machine Learning

אתר הבית של הקורס:

<https://github.com/moshesipper/Applied-Machine-Learning-Course>

Decision and Classification Trees

* עבור כל feature בודקים את מיני ההכרעה
ע"י "Gini Impurity"

Impurity
The level of which a decision is decisive

עבור שני צרכי החלטה אפשריים A, B.

$$\text{Gini-Impurity for leaf} = 1 - P(A)^2 - P(B)^2$$

כאשר נבצע ממוצע ממושקף (כלומר, ביחס למספר

הקואליטיות עבור כל צד) ונקבל את ה - Gini-Impurity

המלא.

במידה וצרכי ההחלטה מש"ס, אנו נבצע ממוצע בין

כל 2 סמוכים. כאשר נוכלם לשאול סל/Yes עבור

כל העל"ם האם הוא קטן או גדול מהערך.

היציג הטוב ביותר שנוכל לקבל הוא 0.

הצד השני ביותר יהיה כמספר הסיווגים: $\frac{1-k}{k}$