

פרויקט סופי ובחינה בעל-פה

בנקים וחברות פיננסיות זקוקים למערכות חיזוי שיכולות לזהות אילו לקוחות צפויים לבצע פעולה פיננסית/עסקה, כדי לשפר התאמת מוצרים, שיווק ממוקד וניהול סיכונים. במסגרת פרויקט הסיום של הקורס, הנכם מתבקשים להתמודד עם האתגר האמור- חיזוי ביצוע עסקה עתידית ללקוחות Santander באמצעות למידת מכונה. תאור האתגר נמצא [כאן](#).

המטרה בפרויקט היא לפתח מודל סיווג בינארי שיחזה האם לקוח יבצע עסקה מסוג מסוים בעתיד ($target=1$) או לא ($target=0$), על בסיס מאפיינים טבלאיים אנונימיים. מטרה משנית היא לספק הסבר לפרדיקציה ספציפית (explainability), ראו [כאן](#).

מתכונת הצגת הפרויקט והבחינה בעל-פה:

1. הפרויקט יבוצע ביחידים.
2. הגשת הפרויקט והצגתו תתבצע באמצעות מחברת ג'ופיטר:
 - א. המחברת תכלול את כל הקוד שכתבתם.
 - ב. המחברת תכלול תאי טקסט להסברים של מה שנעשה, ניתוח התוצאות והמסקנות.
 - ג. גרפים יכללו כותרות ומקרא, ובנוסף הסבר לגרף. כנ"ל טבלאות. יודגש- כל תוצאה היא חסרת חשיבות ללא ניתוח מתאים.
 - ד. יש להגיש את המחברת לתיבת ההגשה המתאימה במודול, עד לבוקר הבחינה.
3. אין להשתמש ב- LLMs למיניהם. עבודות דומות תיפסלנה וינקטו צעדיים משמעתיים.
4. הצגת ההקאתון תתקיים בזום, באופן אינדיבידואלי. חובת הפעלת מצלמה במשך הבחינה. לא ניתן להשתתף בבחינה ללא חיבור מצלמה פעיל כאמור.
5. יש להירשם לבחינה [כאן](#), וזאת עד שבוע לפני הבחינה.
6. במועד הרלוונטי כל סטודנט/ית יציג/תציג את המחברת שלו/ה ואת התוצאות שקיבל/ה. כמובן, מצופה מכל אחד ואחת מכם לשלוט בחומר התאורטי וכמובן בכל רזי הקוד.
7. במסגרת הצגת הפרויקט, יש לפרט של כל מרכיבי האלגוריתם (EDA, עיבוד מקדים, הנדסת מאפיינים, מודל, פירוט הניסויים שבצעתם, תוצאות מרכזיות), תוך התמקדות בעיקר, ובשים לב לזמן שהוקצה לכל הצגה (20 דקות בסך הכל, כולל שאלות).
8. במידה וניתן להעלות את הפרידקציות לאתר Kaggle ולקבל ציון באמצעות האתר, מומלץ לעשות זאת ואת ולהציג את הציון במסגרת הצגת הפרויקט.
9. במהלך ההצגה, אשאל שאלות הנוגעות לפרויקט, אופן ביצועו ושאלות תאורטיות רלוונטיות.

10. הקריטריונים לציון הם כדלקמן :

- ניתוח מעמיק של הבעיה והנתונים, הגדרת מטריקה מתאימה ו- target variable.
- מקוריות ומורכבות הפתרון.
- בחירה נכונה של המודל (לדוגמא, צריכה להיות סיבה טובה למה לא להשתמש במודלים שנחשבים "state of the art").
- הצגת תוצאות, ניתוח תוצאות, הסקת מסקנות מהתוצאות.
- איכות התוצאות. בהקשר זה, יש לבחון אפשרות לשפר את התוצאות באמצעות ביצוע אופטימיזציה של ההיפר-פרמטרים.
- איכות התשובות לשאלות שתשאלו במהלך ההצגה.

לכל שאלה בנושא הפרויקט והבחינה, אנא פנו אליי.

בהצלחה!

דרור