המכללה האקדמית להנדסה סמי שמעון

המכללה האקדמית להנדסה ע"ש סמי שמעון (SCE)

המחלקה להנדסת חשמל ואלקטרוניקה

פרויקט קורס למידת מכונה
"מודל לזיהוי כוס מלאה"

מרצה: דמיטרי בחובסקי

מגישים:
תומר נור מילקוב 325996577
ניקיטה בודובסקי 214156580
אורן
אור

1. מבוא

נושא הפרויקט:

זיהוי מצב מילוי של כוס באמצעות אותות קול מהווה יישום מעניין של למידת מכונה. במהלך מזיגת מים לכוס משתנים המאפיינים האקוסטיים של המערכת עקב שינוי בתהודה (Resonance) ובפיזור התדרים של הקול. שינויים אלו ניתנים למדידה באמצעות מיקרופון ולניתוח באמצעות אלגוריתמים של למידת מכונה.

מטרת פרויקט זה היא לפתח מערכת המסוגלת לזהות האם כוס נמצאת לקראת מצב מלא, בהתבסס על הקול הנוצר במהלך תהליך המילוי בלבד. בניגוד לפתרונות המבוססים על מצלמות, חיישני גובה נוזל או חיישני משקל, המערכת המוצעת עושה שימוש במידע אקוסטי בלבד ולכן מהווה פתרון פשוט, זול ונגיש.

לצורך הפרויקט נאסף מאגר נתונים המכיל הקלטות של מספר סוגי כוסות מחומרים ונפחים שונים. מכל הקלטה הופקו מאפיינים אקוסטיים שונים, ביניהם MFCC, RMS Energy, Spectral Centroid, Spectral Bandwidth, Spectral Rolloff ו־Zero Crossing Rate. מאפיינים אלו שימשו כקלט למודלים שונים של למידת מכונה.

במסגרת העבודה נבחנו מספר אלגוריתמים קלאסיים של למידת מכונה: Logistic Regression, k-NN ו־Random Forest. ביצועי המודלים הוערכו באמצעות Fold-5 Cross Validation וכן באמצעות GroupKFold, במטרה למנוע זליגת מידע בין חלונות שמקורם באותה הקלטה.

תוצאות הניסוי מראות כי ניתן לזהות מאפיינים המעידים על התקרבות הכוס למצב מלא באמצעות ניתוח אודיו בלבד. בנוסף נמצא כי מאפייני תדר, ובפרט מקדמי MFCC, מהווים את הגורם המשמעותי ביותר בקבלת ההחלטה של המודל.

2. איסוף נתונים

(שלב ההקלטות)

לצורך פיתוח המערכת נאסף מאגר נתונים ייעודי הכולל הקלטות אודיו של תהליך מילוי כוסות במים. מטרת איסוף הנתונים הייתה ליצור מגוון מספק של דוגמאות שיאפשר למודל ללמוד את השינויים האקוסטיים המתרחשים במהלך מילוי הכוס ולהבחין בין מצבים בהם הכוס עדיין אינה מלאה לבין מצבים בהם היא מתקרבת למצב מלא.

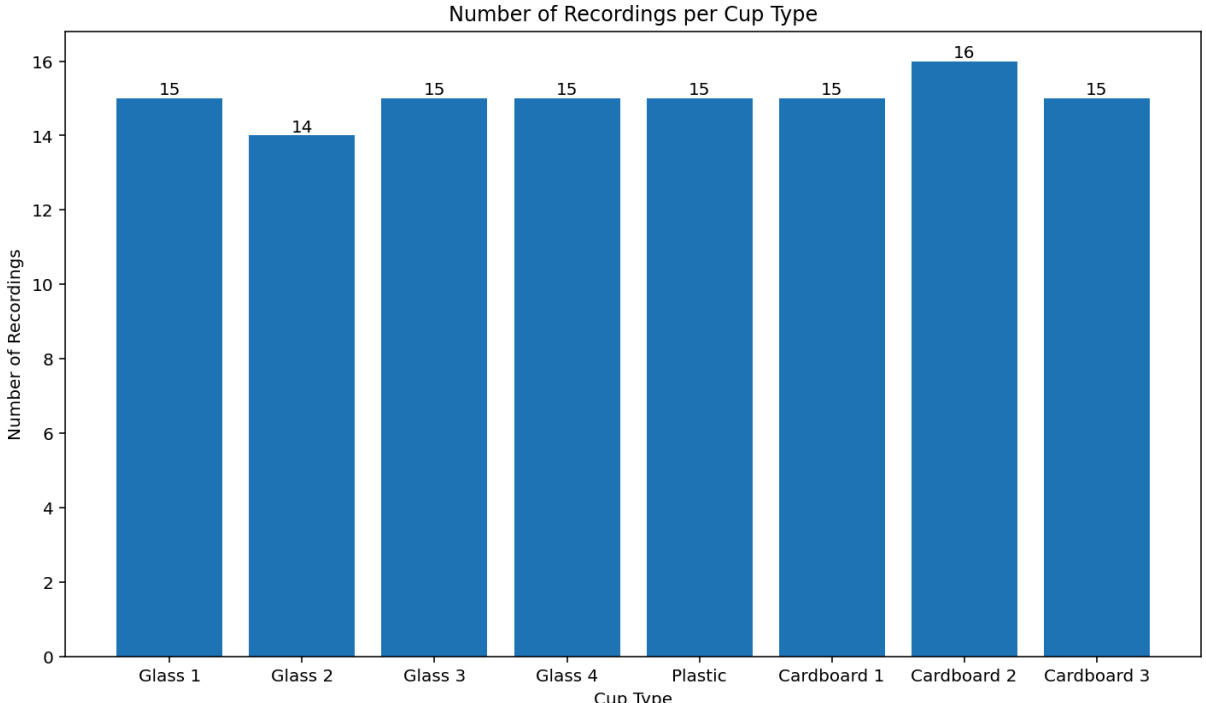
מאגר הנתונים כלל 119 הקלטות שנאספו באופן ידני. ההקלטות בוצעו עבור שמונה סוגי כוסות שונים מחומרים ונפחים מגוונים, ביניהם כוסות זכוכית, פלסטיק וקרטון. עבור כל סוג כוס בוצעו כ-15 הקלטות במטרה ליצור ייצוג מאוזן ככל האפשר של סוגי הכוסות השונים.

במהלך איסוף הנתונים נשמרו מספר תנאים קבועים בהתאם להנחיות הפרויקט. כל הקלטה החלה כאשר הכוס הייתה ריקה לחלוטין, ולאחר מכן בוצעה מזיגת מים עד למילוי הכוס. בנוסף, נשמר מרחק קבוע ככל האפשר בין מקור הקול למיקרופון, וההקלטות בוצעו בסביבה שקטה על מנת לצמצם השפעות של רעשי רקע.

שמירה על תנאי הקלטה אחידים (אותו חדר ואותו ברז) אפשרה להפחית גורמים חיצוניים שעלולים להשפיע על תוצאות המודל, ובכך להבטיח שההבדלים בין ההקלטות ינבעו בעיקר מהשינויים האקוסטיים המתרחשים במהלך תהליך המילוי עצמו.

<u>איור 1</u> מציג את התפלגות ההקלטות בין סוגי הכוסות השונים ששימשו לבניית מאגר הנתונים.

<u>איור 1:</u> (נוצר באמצעות קוד פייתון ומצורף כקובץ dataset_graph.py)



3. אופן ניתוח הנתונים

(בוצע וצורף בקובץ פייתון בשם build_dataset.py)

לאחר איסוף ההקלטות בוצע תהליך עיבוד מקדים (Preprocessing) במטרה להמיר את אותות האודיו למאפיינים מספריים המתאימים ללמידת מכונה.

(תהליך זה מומש בקוד באמצעות ספריית Librosa לעיבוד אודיו וספריית Pandas לניהול הנתונים.)

3.1 חלוקת ההקלטות לחלונות זמן

משך ההקלטות שנאספו לא היה אחיד וכל הקלטה חולקה לחלונות זמן באורך 0.5 שניות. חלוקה זו אפשרה לנתח את השינויים האקוסטיים המתרחשים במהלך תהליך מילוי הכוס ולייצג שלבים שונים בתהליך המילוי באמצעות דוגמאות נפרדות עבור מודל למידת המכונה.

לאחר החלוקה לחלונות התקבל מאגר נתונים המכיל 1495 דוגמאות (Samples), כאשר כל דוגמה מייצגת מקטע אודיו באורך של חצי שנייה.

(בקוד, כל קובץ אודיו נטען באמצעות הפונקציה librosa.load(), ולאחר מכן חולק למקטעים באורך קבוע באמצעות חישוב מספר הדגימות המתאים לחצי שנייה. כל מקטע נשמר בנפרד ועבר תהליך חילוץ מאפיינים.)

3.2 חילוץ מאפיינים (Feature Extraction)

מכל חלון אודיו חולצו מספר מאפיינים המתארים את מאפייני התדר והאנרגיה של הצליל.

(בקוד, תהליך זה בוצע באמצעות פונקציות ייעודיות של ספריית Librosa, כאשר עבור כל חלון חושבו המאפיינים ונשמרו במבנה נתונים אחד.)

MFCC

MFCC (Mel Frequency Cepstral Coefficients) הם מאפיינים נפוצים בתחום עיבוד האודיו והדיבור. מאפיינים אלו מתארים את התפלגות התדרים באות ומאפשרים לזהות שינויים בתכונות האקוסטיות של הצליל.

(עבור כל חלון חולצו 13 מקדמי MFCC באמצעות הפונקציה librosa.feature.mfcc(). לאחר מכן חושבו הממוצע וסטיית התקן של כל אחד מהמקדמים, וערכים אלו נוספו למערך המאפיינים של הדוגמה.)

RMS Energy

מאפיין זה מתאר את עוצמת האות הממוצעת בחלון הנתון. ערך גבוה יותר מעיד בדרך כלל על צליל חזק יותר.

(בקוד חושב מאפיין זה באמצעות הפונקציה librosa.feature.rms(), ומתוכו נשמרו ערכי הממוצע וסטיית התקן.)

Spectral Centroid

מאפיין זה מתאר את "מרכז הכובד" של ספקטרום התדרים. ערך גבוה מצביע על כך שרוב האנרגיה מרוכזת בתדרים גבוהים יותר.

(בקוד חושב באמצעות librosa.feature.spectral_centroid(), ולאחר מכן חושבו עבורו ערכי ממוצע וסטיית תקן.)

Spectral Bandwidth

מאפיין זה מודד את פיזור האנרגיה סביב מרכז הספקטרום ומספק מידע על רוחב תחום התדרים הפעיל.

(בקוד חושב באמצעות librosa.feature.spectral_bandwidth(), ומתוכו נשמרו ערכי הסיכום הסטטיסטיים.)

Spectral Rolloff

מאפיין זה מציין את התדר שמתחתיו נמצאת רוב אנרגיית האות. הוא מאפשר להעריך כיצד האנרגיה מתפלגת בין תדרים נמוכים וגבוהים.

(בקוד חושב באמצעות librosa.feature.spectral_rolloff(), ולאחר מכן חושבו הממוצע וסטיית התקן של הערכים שהתקבלו.)

Zero Crossing Rate

מאפיין זה מודד את מספר המעברים של האות דרך ערך האפס. ערך זה מספק מידע נוסף על אופי האות ותכולת התדרים שלו.

(בקוד חושב באמצעות librosa.feature.zero_crossing_rate(), ומתוכו נשמרו ערכי הממוצע וסטיית התקן כחלק ממערך המאפיינים.)

3.3 יצירת מערך הנתונים

לאחר חילוץ כל המאפיינים נבנה מערך נתונים (Dataset) שבו כל שורה מייצגת חלון אודיו אחד, וכל עמודה מייצגת מאפיין שחולץ מהאות.

בנוסף לכל חלון הוצמד תיוג (Label) המתאר האם החלון שייך לשלב שבו הכוס עדיין אינה מלאה או לשלב שבו הכוס נמצאת לקראת מצב מלא. תיוג זה שימש כמשתנה המטרה (Target Variable) עבור מודלי למידת המכונה.

מערך הנתונים הסופי שימש לאימון ולהערכת ביצועי המודלים השונים שנבחנו במסגרת הפרויקט.

4. איך המערכת מגיעה להחלטה

לאחר יצירת מערך הנתונים וביצוע חילוץ המאפיינים, נבחנו מספר מודלים קלאסיים של למידת מכונה במטרה לזהות האם חלון אודיו מסוים שייך לשלב שבו הכוס עדיין אינה מלאה או לשלב שבו היא מתקרבת למצב מלא.

4.1 מודלים שנבדקו

במסגרת הפרויקט נבחנו מספר אלגוריתמים נפוצים של למידת מכונה:

Logistic Regression

מודל לינארי המשמש לבעיות סיווג. המודל מחשב את ההסתברות שכל דוגמה שייכת לאחת מהמחלקות ומבצע סיווג בהתאם.

(k-Nearest Neighbors) k-NN

מודל המבצע סיווג על סמך הדוגמאות הקרובות ביותר במרחב המאפיינים. עבור כל דוגמה חדשה נבדקים השכנים הקרובים ביותר, והסיווג נקבע לפי הרוב ביניהם.

Random Forest

מודל המבוסס על אוסף גדול של עצי החלטה. כל עץ מקבל החלטה באופן עצמאי, והחלטת המודל הסופית מתקבלת באמצעות הצבעת רוב בין כל העצים. יתרון משמעותי של מודל זה הוא האפשרות להעריך את חשיבות המאפיינים השונים בתהליך קבלת ההחלטה.

4.2 הערכת ביצועים

בהתאם להנחיות הפרויקט נעשה שימוש בשיטת 5-Fold Cross Validation לצורך הערכת ביצועי המודלים.

בשיטה זו מערך הנתונים מחולק לחמישה חלקים. בכל איטרציה ארבעה חלקים משמשים לאימון המודל וחלק אחד משמש לבדיקה. התהליך חוזר חמש פעמים כך שכל חלק משמש פעם אחת כסט בדיקה.

בנוסף נעשה שימוש ב־GroupKFold. שיטה זו מבטיחה שכל החלונות שמקורם באותה הקלטה יופיעו באותו קיפול (Fold), ובכך נמנעת זליגת מידע בין סט האימון לסט הבדיקה.

לצורך ההשוואה חושבו המדדים הבאים:

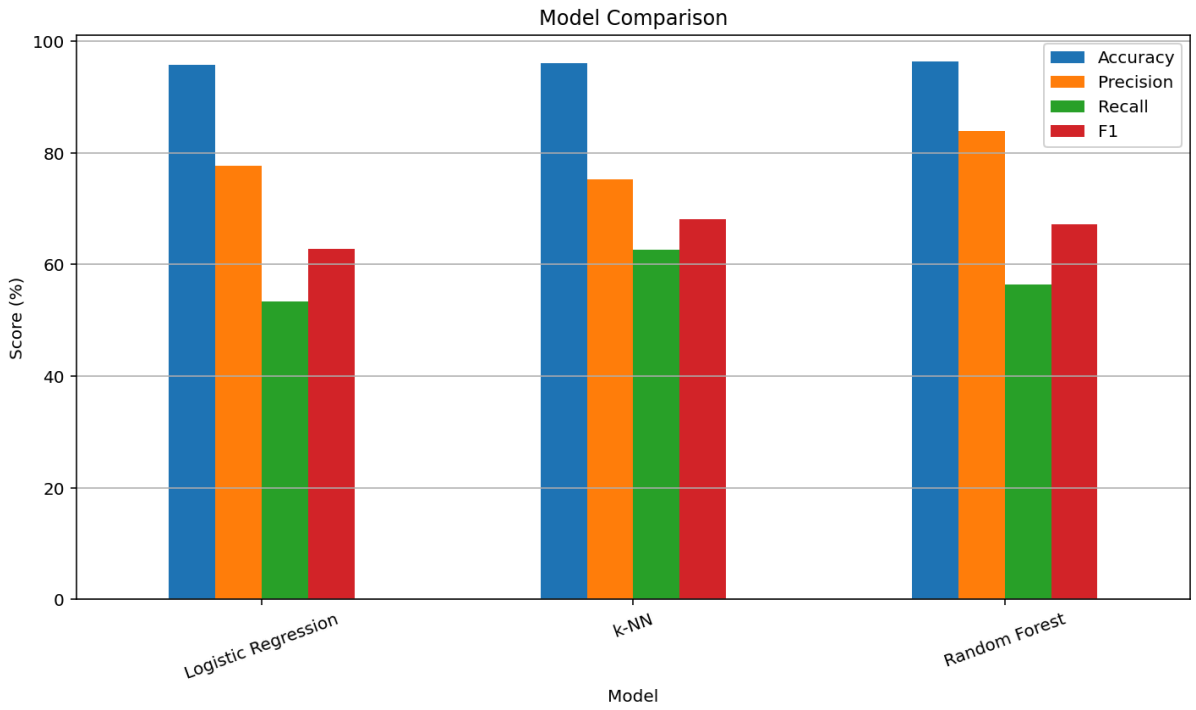
- Accuracy
- Precision
- Recall
- F1 Score

4.3 תוצאות המודלים

<u>איור 2</u> מציג השוואה בין ביצועי המודלים שנבדקו.

(מצורף כקובץ פייתון model_comparison_graph.py)

<u>איור 2:</u>



(טבלה זו נבנתה על ידי פלט הקוד train_model.py המצורף)

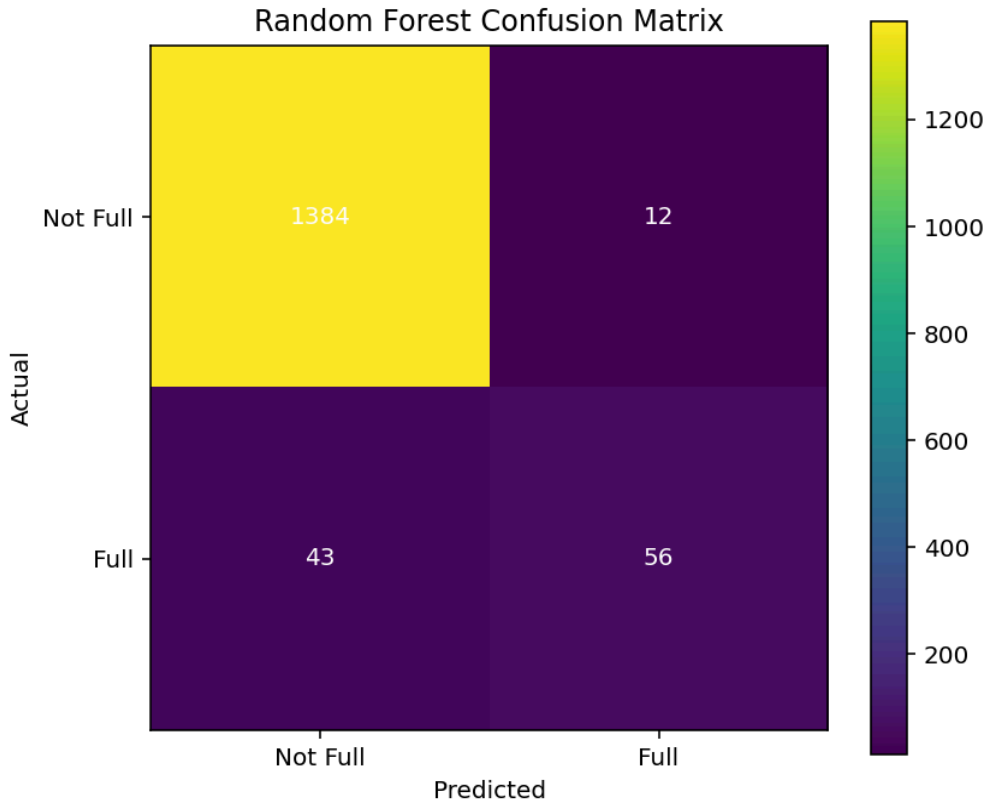
מודל	Accuracy	Precision	Recall	F1
Logistic Regression	95.85%	77.71%	53.37%	62.76%
k-NN	96.12%	75.29%	62.58%	68.07%
Random Forest	96.32%	83.93%	56.47%	67.18%

מבין המודלים שנבדקו, מודל Random Forest השיג את ערך ה־Accuracy הגבוה ביותר וכן את ערך ה-Precision הגבוה ביותר. בנוסף, המודל מאפשר ניתוח של חשיבות המאפיינים השונים ולכן נבחר כמודל הסופי של המערכת.

על מנת להבין טוב יותר את ביצועי המודל, נבחנה מטריצת הבלבול (Confusion Matrix) של מודל Random Forest. (איור 3)

איור 3:

(איור זה נבנה על ידי קוד פייתון המצורף בשם confusion_matrix_graph.py)



ניתן לר
מלא", "לא

בנוסף, מתוך הדוגמאות שסומנו כ"מלא", 56 דוגמאות זוהו בצורה נכונה ו־43 דוגמאות סווגו בטעות כ"לא מלא".

תוצאה זו תואמת את ערכי ה־Accuracy וה־Precision הגבוהים שהתקבלו עבור המודל, ומראה כי המערכת מצליחה לזהות בצורה טובה את רוב המצבים במאגר הנתונים.

4.4 בחירת המודל הסופי

מודל Random Forest נבחר כמודל הסופי של המערכת. המודל השיג את ביצועי הסיווג הטובים ביותר מבין המודלים שנבחנו במסגרת העבודה, וכן מאפשר ניתוח של חשיבות המאפיינים השונים בתהליך קבלת ההחלטה.

הבחירה בוצעה מכיוון שמודל זה השיג ביצועים גבוהים מאוד, מאפשר ניתוח של חשיבות המאפיינים השונים, ונלמד במסגרת הקורס. בנוסף, מודל Random Forest נחשב למודל יציב וקל יחסית להסבר ולניתוח.

קובץ רלוונטי בפרויקט: train_model.py - אימון המודלים, ביצוע 5-Fold Cross Validation והשוואת הביצועים.

5. ניתוח החלטת המודל (Feature Importance)

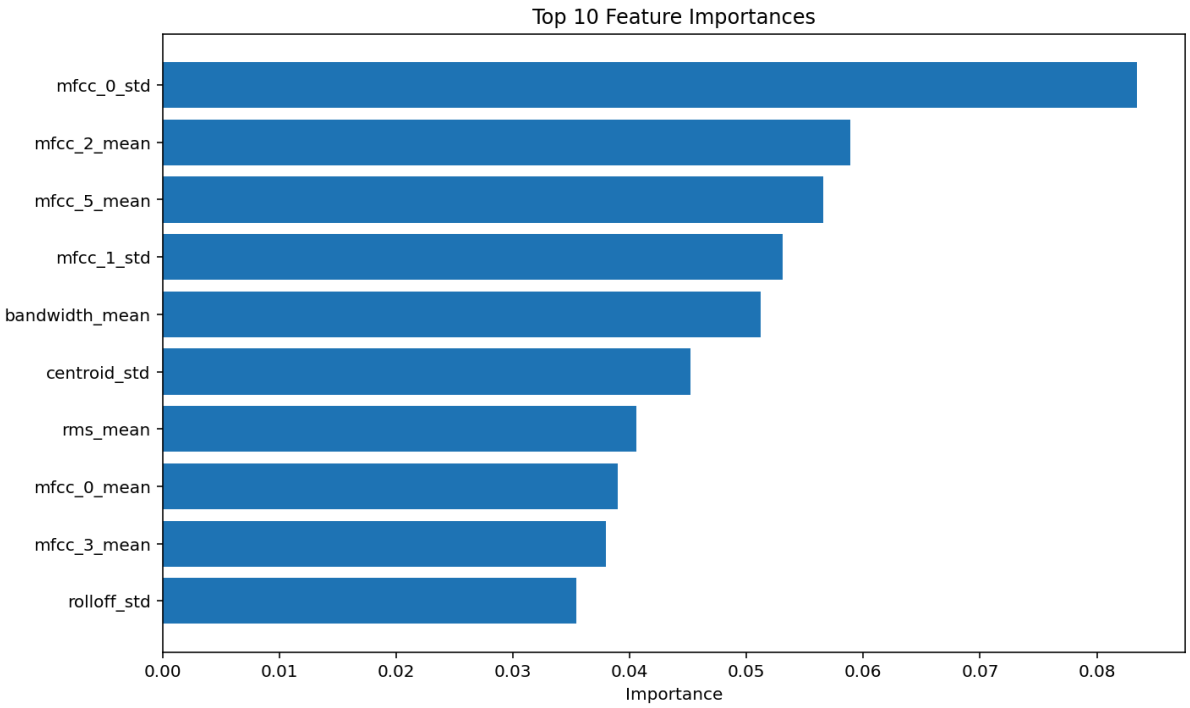
אחד היתרונות המרכזיים של מודל Random Forest הוא היכולת להעריך את חשיבות המאפיינים השונים בתהליך קבלת ההחלטה. יכולת זו מאפשרת להבין אילו מאפיינים תרמו בצורה המשמעותית ביותר לביצוע הסיווג.

לאחר אימון המודל חושבה חשיבות המאפיינים (Feature Importance) עבור כלל המאפיינים שחולצו מהאות. איור 4 מציג את עשרת המאפיינים בעלי התרומה הגבוהה ביותר.

איור 4:

(איור זה נוצר באמצעות קוד פייתון המצורף בשם feature_importance_graph.py)

עשרת המאפיינים בעלי התרומה הגבוהה ביותר להחלטת מודל Random Forest.



ניתן לראות כי רוב המאפיינים בעלי החשיבות הגבוהה ביותר שייכים למשפחת מאפייני ה־MFCC. בפרט, המאפיינים mfcc_0_std, mfcc_2_mean ו־mfcc_5_mean קיבלו את ציוני החשיבות הגבוהים ביותר מבין כלל המאפיינים שנבחנו.

מאפייני MFCC מתארים את מבנה התדרים של הצליל ולא את עוצמתו בלבד. תוצאה זו מעידה כי המודל אינו מקבל החלטות על סמך עוצמת זרם המים בלבד, אלא בעיקר על סמך השינויים האקוסטיים המתרחשים במהלך מילוי הכוס.

ככל שמפלס המים עולה, משתנים מאפייני התהודה (Resonance) של הכוס, דבר הגורם לשינוי בהתפלגות התדרים של הצליל המוקלט. המודל למד לזהות שינויים אלו ולהשתמש בהם לצורך קביעת מצב המילוי של הכוס.

בנוסף למאפייני ה־MFCC נמצאו גם מאפיינים נוספים בעלי חשיבות, כגון Spectral Bandwidth, Spectral Centroid ו־RMS Energy. מאפיינים אלו מספקים מידע נוסף על התפלגות האנרגיה והתדרים באות.

ממצאים אלו מחזקים את ההשערה שניתן לזהות את מצב מילוי הכוס באמצעות ניתוח מאפייני התדר של הצליל בלבד, ללא צורך בחיישנים נוספים.

6. סיכום, תובנות והצעות לשיפור

בפרויקט זה פותחה מערכת המבוססת על למידת מכונה שמטרתה לזהות האם כוס נמצאת לקראת מצב מלא באמצעות ניתוח אותות אודיו בלבד. לצורך כך נאסף מאגר נתונים הכולל 119 הקלטות של תהליך מילוי כוסות מסוגים שונים, אשר חולקו לחלונות זמן ועברו תהליך חילוץ מאפיינים.

במהלך העבודה נבחנו מספר מודלים של למידת מכונה, ביניהם k-NN, Logistic Regression, ו־Random Forest. ביצועי המודלים הוערכו באמצעות Fold-5 Cross Validation תוך שימוש ב־GroupKFold למניעת זליגת מידע בין דוגמאות מאותה הקלטה.

תוצאות הניסוי הראו כי ניתן לזהות את מצב המילוי של הכוס באמצעות מידע אקוסטי בלבד. מבין המודלים שנבחנו, מודל Random Forest נבחר כמודל הסופי בזכות ביצועיו הגבוהים ויכולתו לספק מידע על חשיבות המאפיינים השונים בתהליך קבלת ההחלטה.

ניתוח חשיבות המאפיינים הראה כי מאפייני MFCC היו בעלי התרומה הגבוהה ביותר לביצועי המודל. ממצא זה מצביע על כך שהמודל למד לזהות שינויים במבנה התדרים של הצליל הנגרמים כתוצאה משינוי מפלס המים בכוס.

במהלך הפרויקט התמודדנו עם מספר אתגרים, ביניהם איסוף נתונים עקבי, שמירה על תנאי הקלטה דומים ותהליך תיוג הנתונים. למרות מגבלות אלו, המודל הצליח להגיע לביצועי סיווג גבוהים ולהבחין בין שלבי המילוי השונים.

בעתיד ניתן לשפר את המערכת באמצעות הגדלת מאגר הנתונים, הוספת סוגי כוסות נוספים, ביצוע תיוג מדויק יותר של רגע המילוי וכן בחינת ביצועי המערכת בסביבות רועשות ובתנאי הקלטה מגוונים יותר.

לסיכום, תוצאות העבודה מדגימות כי שילוב של עיבוד אותות ולמידת מכונה מאפשר לזהות את מצב מילוי הכוס בצורה יעילה, ללא צורך בחיישנים פיזיים נוספים, תוך הסתמכות על מידע אקוסטי בלבד.