Progress Report-I

Surface type detection for the robot's indoor navigation using Signal Processing and Machine Learning

Guided by Dr Seethalekshmi K

Team

Akansh Maurya, Akash Singh. Lokendra Rathore

Objective for this stage

1. Study of signal/data, perform a description of data, correlation of various signal(Pearson correlation).
2. Compare the accuracy of various ML algorithm for prediction of the floor on original data.
3. Compare the accuracy of various ML algorithm for prediction of the floor on denoised data(Fourier and wavelet denoising).

1. Data description

- X_[train].csv - the input data, covering 10 sensor channels and 128 measurements per time series plus three ID columns:
    - row_id: The ID for this row.
    - series_id: ID number for the measurement series.
    - measurement_number: Measurement number within the series.
    - Sensor channels:
        - Sensor Channel 0:orientation_X
        - Sensor Channel 1:orientation_Y
        - Sensor Channel 2:orientation_Z
        - Sensor Channel 3:orientation_W
        - Sensor Channel 4:angular_velocity_X
        - Sensor Channel 5:angular_velocity_Y
        - Sensor Channel 6:angular_velocity_Z
        - Sensor Channel 7:linear_acceleration_X
        - Sensor Channel 8:linear_acceleration_Y

        - Sensor Channel 9:linear_acceleration_Z
- Y_[target].csv: the target, the prediction vector.
    - series_id: ID number for the measurement series.
    - group_id: ID number for all of the measurements taken in a recording session. Provided for the training set only, to enable more cross-validation strategies.
    - surface: the target.

	series_id	group_id	surface
0	0	13	fine_concrete
1	1	31	concrete
2	2	20	concrete
3	3	31	concrete
4	4	22	soft_tiles
5	5	1	tiled
6	6	34	soft_pvc
7	7	31	concrete
8	8	33	hard_tiles_large_space
9	9	11	tiled

Figure 1: A look into target

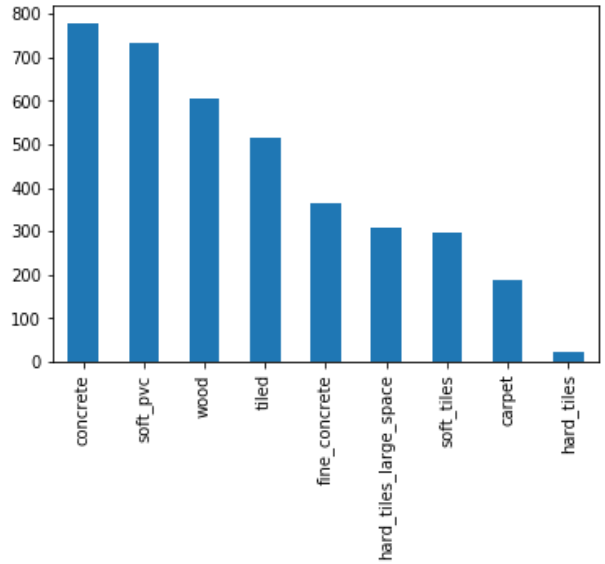
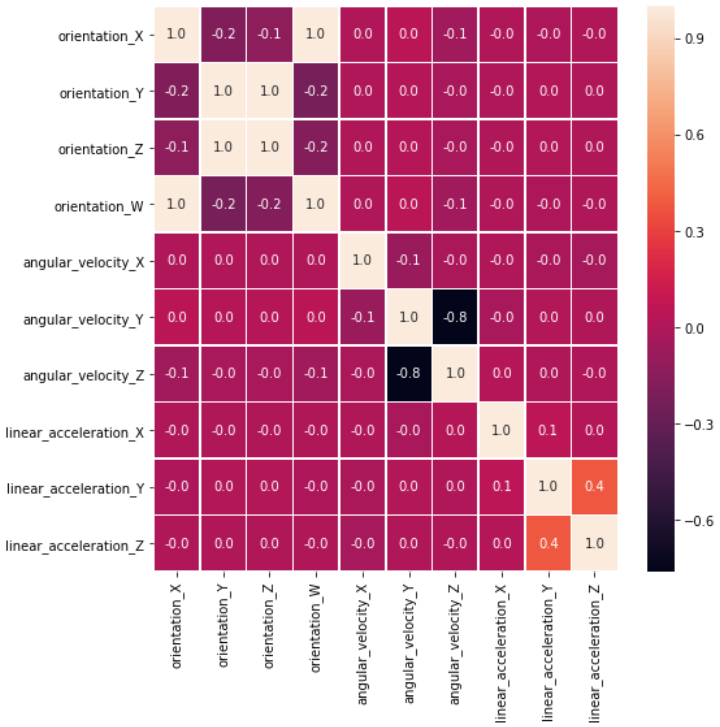


Figure 2: Distribution of classes

Well, this is important, there is a strong correlation between:

- angular_velocity_Z and angular_velocity_Y
- orientation_X and orientation_Y
- orientation_Y and orientation_Z

We can perform experiments in which we will remove one of the correlated features in order to save computation and complexity of the model.



More detailed explanation:
https://github.com/akansh12/help_robots_help_humanity/blob/main/EDA.ipynb

2. Compare the accuracy of different ML algorithm on original data.

We used the following statistic features:

1. Mean
2. Standard Deviation
3. Max
4. Min
5. Max to Min ratio
6. First-order differentiation
7. Second-order differentiation

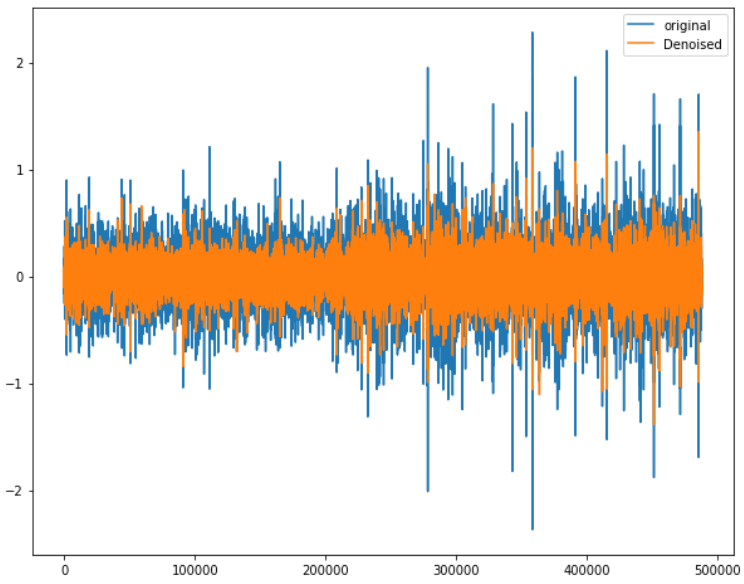
We received a baseline accuracy of **47.68381%** on the test dataset.

	MLA Name	MLA Parameters	MLA Train Accuracy	MLA Train Accuracy Mean	MLA Test Accuracy	MLA Test Accuracy Mean	MLA Test Accuracy Std
3	GradientBoostingClassifier	{'ccp_alpha': 0.0, 'criterion': 'friedman_mse'...	[0.9973761889143982, 0.9908136482939632, 0.996...	0.993635	[0.48751642575558474, 0.37139107611548555, 0.4...	0.476381	0.0558823
20	XGBClassifier	{'objective': 'binary:logistic', 'use_label_en...	[1.0, 1.0, 1.0, 1.0, 1.0]	1	[0.49934296977660975, 0.3727034120734908 4, 0.4...	0.471669	0.0576054
4	RandomForestClassifier	{'bootstrap': True, 'ccp_alpha': 0.0, 'class_w...	[1.0, 1.0, 1.0, 1.0, 1.0]	1	[0.4980289093298292, 0.34908136482939633, 0.42...	0.469048	0.0722208
2	ExtraTreesClassifier	{'bootstrap': False, 'ccp_alpha': 0.0, 'class_...	[1.0, 1.0, 1.0, 1.0, 1.0]	1	[0.5256241787122208, 0.3320209973753281, 0.424...	0.460394	0.0749429
1	BaggingClassifier	{'base_estimator': None, 'bootstrap': True, 'b...	[0.9931124959002952, 0.9957349081364829, 0.994...	0.993635	[0.507227332457293, 0.30708661417322836, 0.385...	0.431791	0.0754181
16	DecisionTreeClassifier	{'ccp_alpha': 0.0, 'class_weight': None, 'crit...	[1.0, 1.0, 1.0, 1.0, 1.0]	1	[0.43626806833114323, 0.3320209973753281, 0.39...	0.407885	0.0426353

3. Compare the accuracy of different ML algorithm on Denoised Data (FFT).

We used the same features just performed denoising using FFT.

We are performing in order to improve the accuracy, alas it didn't happen and we noticed a decrease in the accuracy and increase in Standard deviation.



	MLA Name	MLA Parameters	MLA Train Accuracy	MLA Train Accuracy Mean	MLA Test Accuracy	MLA Test Accuracy Mean	MLA Test Accuracy Std
2	ExtraTreesClassifier	{'bootstrap': False, 'ccp_alpha': 0.0, 'class_...	[1.0, 1.0, 1.0, 1.0, 1.0]	1	[0.5282522996057819, 0.2979002624671916, 0.428...	0.457769	0.0885261
3	GradientBoostingClassifier	{'ccp_alpha': 0.0, 'criterion': 'friedman_mse'...	[0.9970482125286979, 0.9950787401574803, 0.995...	0.995079	[0.4980289093298292, 0.3293963254593176, 0.457...	0.454341	0.068532
20	XGBClassifier	{'objective': 'binary:logistic', 'use_label_en...	[1.0, 1.0, 1.0, 1.0, 1.0]	1	[0.4980289093298292, 0.33858267716535434, 0.42...	0.451987	0.0638282
4	RandomForestClassifier	{'bootstrap': True, 'ccp_alpha': 0.0, 'class_w...	[1.0, 1.0, 1.0, 1.0, 1.0]	1	[0.5006570302233903, 0.2874015748031496, 0.420...	0.448052	0.0883982
1	BaggingClassifier	{'base_estimator': None, 'bootstrap': True, 'b...	[0.9950803542144966, 0.994750656167979, 0.9940...	0.994751	[0.5111695137976346, 0.2992125984251969, 0.394...	0.435726	0.0800442
16	DecisionTreeClassifier	{'ccp_alpha': 0.0, 'class_weight': None, 'crit...	[1.0, 1.0, 1.0, 1.0, 1.0]	1	[0.4244415243101183, 0.2874015748031496, 0.382...	0.396336	0.0645472

Next Task

1. Perform Wavelet denoising methods
2. Perform advance feature selection (Frequency domain) to boost the accuracy of ML algorithms.
3. Deep Learning Methods for the predictions.