# impurity-based feature importance

Below the plot of feature importance from Random Forest fitted on 96 price series, volume and dummies. Features here are elements of vector, flattened matrix [variables, t].

Obraz zawierający transport, statek wodny, żaglowiec, maszt

Opis wygenerowany automatycznie

Below list of prices indices sorted from the most to the least important (trajectories were summed). We can see that indices close to traded delivery (59) are the most important. For now no idea why ~40 are so important too.

np.argsort(np.sum(importances[:-1101].reshape((96,1100)),1))[::-1]

array([58, 59, 56, 62, 57, 36, 53, 54, 29, 35, 40, 43, 70, 55, 52, 92, 30,

63, 7, 39, 61, 42, 49, 44, 38, 25, 23, 24, 41, 18, 31, 5, 60, 76,

50, 78, 51, 64, 87, 32, 48, 67, 13, 6, 9, 3, 14, 16, 17, 28, 11,

10, 33, 90, 12, 45, 21, 46, 69, 2, 27, 34, 22, 0, 4, 37, 26, 15,

71, 94, 68, 1, 79, 8, 66, 86, 19, 47, 80, 74, 20, 75, 88, 95, 77,

81, 72, 83, 73, 65, 84, 85, 82, 91, 93, 89], dtype=int64)