Dokumentacja

Metody w konstruktorze przyjmują dwie wartości:

**:param** *y: szereg czasowy, dla którego dopasowany zostanie model. Szereg powinien mieć indeks miesięczny (MS - month start)***:param** *horizon: wartość większa od 0 lub None. W przypadku wartości liczbowej oznacza to, z jakich danych może co najwyżej korzystać model podczas wykonywania predykcji. Przykładowo, obliczając predykcję na marzec 2020 z horyzontem 1 miesiąc, model będzie wykorzystywał dane do lutego 2020. Jednak w przypadku horyzontu 2, model wykorzysta co najwyżej dane ze stycznia 2021. Wyjątkiem jest jedynie sytuacja, gdy dane te zostają przekazane jako szereg treningowy w konstruktorze.*

*Przykładowo, jako dane treningowe wykorzystano dane 01.01.2017-01.05.2019. Wykonując teraz predykcję dla 01.06.2019 z horyzontem 2 zostaną wykorzystane dane do 01.05.2019, dla 01.07.2019, także zostaną wykorzystane dane do 01.05.2019, natomiast dla 01.08.2019 zostaną już wykorzystane dane do 01.06.2019 (o ile zostały one przekazane w parametrach metody predict). Uzasadnieniem takiego postępowania jest założenie, że dane treningowe powinny stanowić minimalny zasób danych dostępnych dla modelu, a nowe napływające dane mogą go co najwyżej uzupełniać.*

*W przypadku wartości None, model wykorzystuje jedynie dane treningowe do wykonywania predykcji.*

* plot\_acf oraz plot\_pacf

Obie powyższe metody rysują odpowiednie wykresy dla danych treningowych przekazanych w konstruktorze

* hiperparameter\_search\_fit

Przeszukuje wszystkie kombinacje parametrów p, d, q, P, D, Q z wybranym s oraz wybiera najlepsze hiperparametry w kontekście przekazanej metryki (dla której mamy minimalne wartości metryki). W przypadku mse (mean squared error) konieczne jest wydzielenie zbioru walidacyjnego. W tym celu wykorzystywany jest parametr „split\_fraction”. Przykładowo, dla domyślnej wartości 0.8, pierwsze 80% obserwacji zostanie przydzielonych do zbioru treningowego, natomiast pozostałe 20% do zbioru walidacyjnego na którym będzie szacowany błąd modelu.

Wybrane parametry służą do budowania modelu. Nie jest konieczne wywoływanie metody fit.

* fit

Drugi sposób po hiperparameter\_search\_fit na ustawienie odpowiednich parametrów aby móc wykonywać predykcje.

* predict

Metoda predict w dużej mierze wykorzystuje parametr test\_data. Jest to pd.Series z indeksem, dla którego zostanie wygenerowana predykcja. Szereg może zawierać NaNy. Jest to przydatne w sytuacji, kiedy chcemy obliczyć predykcję dla miesięcy przyszłych, dla których nie ma jeszcze danych. Dane tego szeregu wykorzystywane są jedynie w przypadku horizon ustawionego na pewną wartość liczbową w konstruktorze (czyli inną niż None). W takiej sytuacji wykonując predykcje dla pewnego momentu, algorytm będzie próbował uzupełnić dane treningowe o test\_data, które poprzedzają bieżący moment o co najmniej tyle miesięcy ile wynosi horizon. Za każdym razem tworzony będzie nowy model, jednak wszystkie te modele zachowają te same parametry order oraz seasonal\_order.

**Warto pamiętać, że zbiór treningowy nie jest nigdy ograniczany. Czyli wykonując predykcję na luty 2020, jeśli w danych treningowych był dostępny styczeń 2020 to zostanie on wykorzystany do wykonania predykcji nawet jeśli horyzont jest większy niż 1.**

W przypadku ustawienia plot=True, rysowany jest wykres porównujący predykcję wraz z przedziałami ufności do rzeczywistych wartości.

Metoda zwraca predykcję w formie pd.Series z indeksem jednakowym jak test\_data

* analyse\_results

Zwraca różne przydatne informacje na temat modelu wytrenowanego na zbiorze treningowym przekazanym w konstruktorze. Wyświetlany jest wykres diagnostyczny oraz wypisywane podsumowanie w terminalu.