Modele w przestrzeni stanów. Ocena jakości i porównanie modeli.

Model w przestrzeni stanów to abstrakcyjny matematyczny model, który służy do opisywania systemów np. dynamicznych, w których stan może być opisany przez zbiór zmiennych stanu.

**Model w przestrzeni stanów składa się z trzech elementów:**

* Zbiór stanów - określa wszystkie możliwe stany, w jakich może znaleźć się system. Zbiór stanów może być ograniczony lub nieograniczony, w zależności od tego, czy istnieje ograniczenie na liczbę możliwych stanów systemu.

* Zbiór symboli wejściowych - określa wszystkie możliwe wejścia, które system może otrzymać. Zbiór symboli wejściowych może być ograniczony lub nieograniczony, w zależności od tego, czy istnieje ograniczenie na liczbę możliwych symboli wejściowych.

* Funkcja przejścia - określa, jak system przechodzi z jednego stanu do innego w odpowiedzi na wejście. Funkcja przejścia może być opisana za pomocą tabeli lub funkcji matematycznej, która przypisuje każdemu stanowi i wejściu odpowiedni nowy stan.

Model w przestrzeni stanów może również zawierać dodatkowe elementy, takie jak zbiór symboli wyjściowych, który określa wszystkie możliwe wyjścia systemu, oraz funkcję wyjścia, która określa, jakie wyjście jest generowane w danym stanie.

Model w przestrzeni stanów jest często używany do opisywania systemów cyfrowych, takich jak komputery, ale może być również stosowany do opisywania innych typów systemów, takich jak systemy kontroli procesów przemysłowych czy układy zautomatyzowanego transportu.

Kategorie modeli:

* Liniowe vs Nieliniowe
* Statyczne vs Dynamiczne
* Deterministyczne vs Niedeterministyczne
* Skupione vs Rozproszone

Model liniowy - W swojej najprostszej postaci model liniowy określa (liniowe) powiązanie pomiędzy zmienną zależną (lub odpowiedzią) Y, a zbiorem predykatorów (zmiennych objaśniających) X. Dzięki takiemu modelowi można na przykład ocenić (przewidzieć) wagę człowieka na podstawie jego wzrostu i płci.

Model nieliniowy - model o nieliniowych parametrach (hehe) czyli parametry nie są stałe, a zmienne. Ogólnie nieliniowość to cecha układu polegająca na tym, że wartość wyjściowa nie jest wprost proporcjonalna do danych wejściowych.

Model statyczny - stan modelu w danej chwili t jest zależny wyłącznie od wartości parametrów w chwili t.

Model dynamiczny - stan modelu w danej chwili t jest zależny nie tylko od wartości parametrów w chwili t, ale także od parametrów z przeszłości.

Model deterministyczny jest modelem, w którym każde przejście jest jednoznacznie określone przez aktualny stan i przysługujące mu wejście. Oznacza to, że dla danego stanu i wejścia, model zawsze będzie przechodził w ten sam stan wyjściowy.

Model niedeterministyczny natomiast jest modelem, w którym dla danego stanu i wejścia, mogą wystąpić różne stany wyjściowe. Oznacza to, że dla tego samego stanu i wejścia, model może przejść do różnych stanów wyjściowych.

Model skupiony - umożliwia kategoryzowanie rekordów w określoną liczbę klastrów. Może to pomóc w zidentyfikowaniu naturalnych grup w danych. Modele grupowania koncentrują się na identyfikacji grup podobnych rekordów i oznaczaniu rekordów zgodnie z grupą, do której należą.

Model rozproszony - nie znalazłem żadnej sensownej definicji, ale wydaje się naturalną przeciwnością modelu skupionego.

# Ocena jakości i porównanie modeli

Istnieje wiele metod oceny jakości modelu, które można podzielić na dwie główne kategorie: metody bezwzględne i względne.

Metody bezwzględne polegają na porównaniu przewidywań modelu z rzeczywistymi danymi. Przykładami takich metod są:

* Błąd średniokwadratowy (MSE) - mierzy średni kwadrat różnicy między przewidywaniami a rzeczywistymi danymi.
* Błąd bezwzględny średni (MAE) - mierzy średnią różnicę między przewidywaniami a rzeczywistymi danymi.
* Błąd względny średni (MAPE) - mierzy średnią procentową różnicę między przewidywaniami a rzeczywistymi danymi.
* Błąd bezwzględny maksymalny (MAXAE) - mierzy maksymalną różnicę między przewidywaniami a rzeczywistymi danymi.
* Krzywa ROC - jest to wykres pokazujący zależność między częstotliwością fałszywie pozytywnych i prawdziwie pozytywnych wyników, przydatne do oceny modeli klasyfikacyjnych.
* Współczynnik podobieństwa (SIM) - mierzy stopień podobieństwa między przewidywaniami a rzeczywistymi danymi.
* Testy statystyczne - takie jak test chi-kwadrat, które pozwalają na porównanie jakości modelu z innymi modelami.
* Współczynnik determinacji (R^2) - mierzy procentową zmienność rzeczywistych danych, która jest opisana przez model.

Metody względne polegają na porównaniu przewidywań modelu z przewidywaniami innych modeli. Przykładami takich metod są:

* Porównanie z modelem bazowym - polega na porównaniu przewidywań modelu z przewidywaniami modelu bazowego, np. modelu prostego.
* Porównanie z modelem konkurencyjnym - polega na porównaniu przewidywań modelu z przewidywaniami innego modelu, który jest uważany za konkurencyjny.
* Test Akaike'a - jest to metoda statystyczna pozwalająca na porównanie jakości modeli na podstawie ich złożoności i dokładności.
* Test Schwarz - jest to metoda statystyczna pozwalająca na porównanie jakości modeli na podstawie ich złożoności i dokładności.
* Test Bayesowski - jest to metoda statystyczna pozwalająca na porównanie jakości modeli na podstawie ich złożoności i dokładności z uwzględnieniem prawdopodobieństwa a priori.

Wybór odpowiedniej metody oceny jakości modelu zależy od kontekstu i celów analizy.

Warto pamiętać, że te metody porównawcze mają swoje ograniczenia i są one przydatne tylko dla porównywania modeli z pewnymi założeniami, np. założenie normalności danych, równości wariancji itp. Dlatego, zalecane jest, aby wykorzystać kilka metod porównawczych, by uzyskać bardziej dokładny obraz jakości modelu.