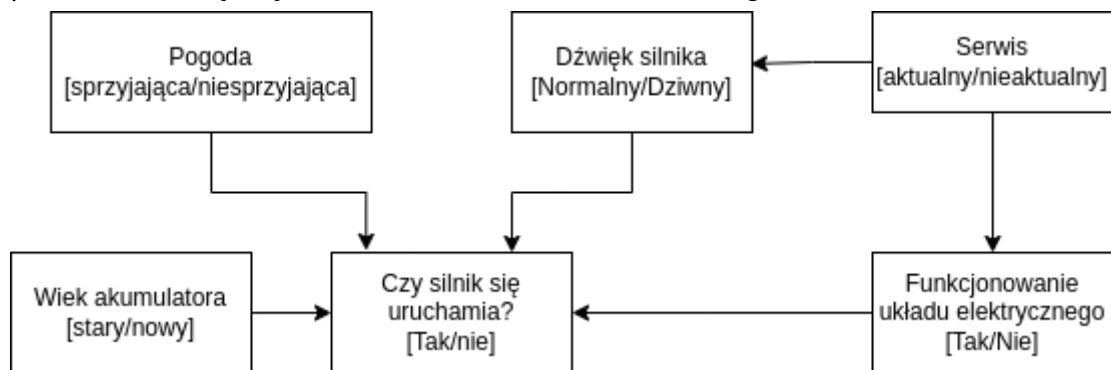


WSI 23Z LAB7 - Sieci Bayesowskie

Maksym Bieńkowski

Szczegóły implementacyjne

Zaimplementowana przeze mnie prosta sieć bayesowska ma na celu umożliwienie diagnostyki problemów związanych z silnikiem samochodu. Oto diagram zależności w sieci:



Tabele prawdopodobieństw warunkowych zachodzących między poszczególnymi wierzchołkami sieci wyznaczone zostały heurystycznie i zostały umieszczone w pliku `src/probs.py`.

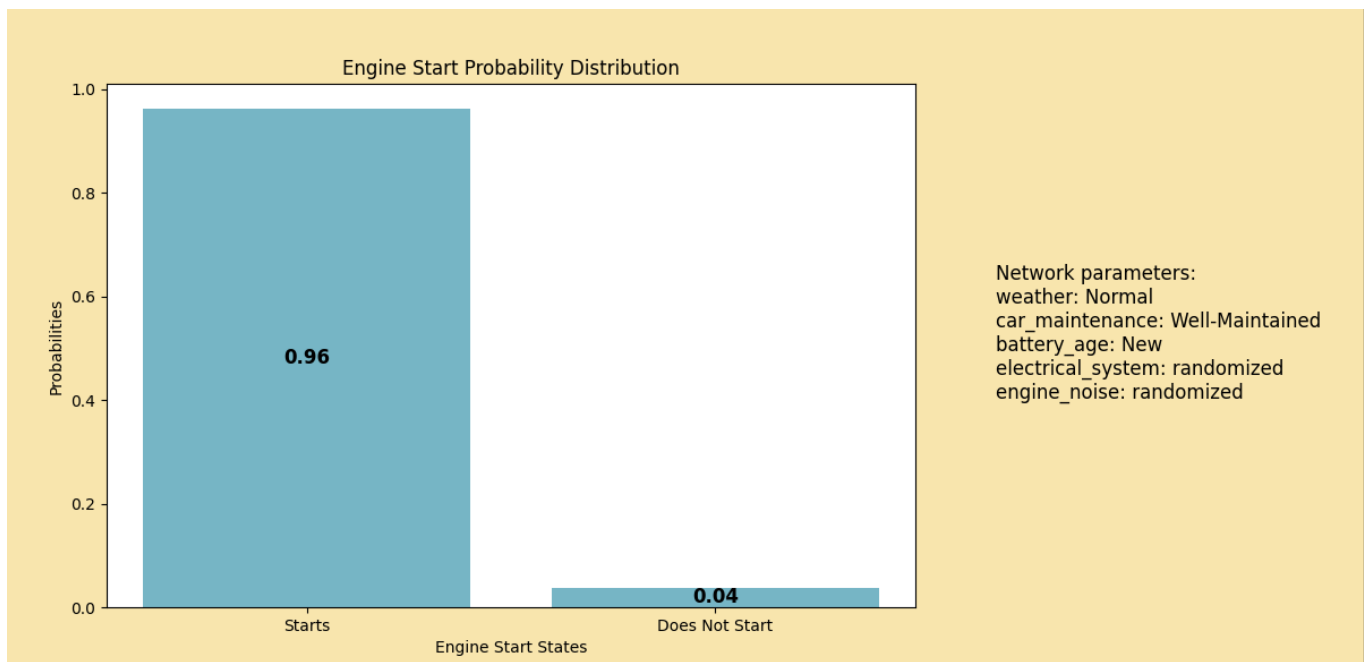
Postanowiłem także eksperymentalnie wykazać działanie niektórych praw warunkowej niezależności w sieci.

Eksperymenty i wyniki

Badanie prawdopodobieństw udanego uruchomienia silnika na podstawie różnych danych wejściowych

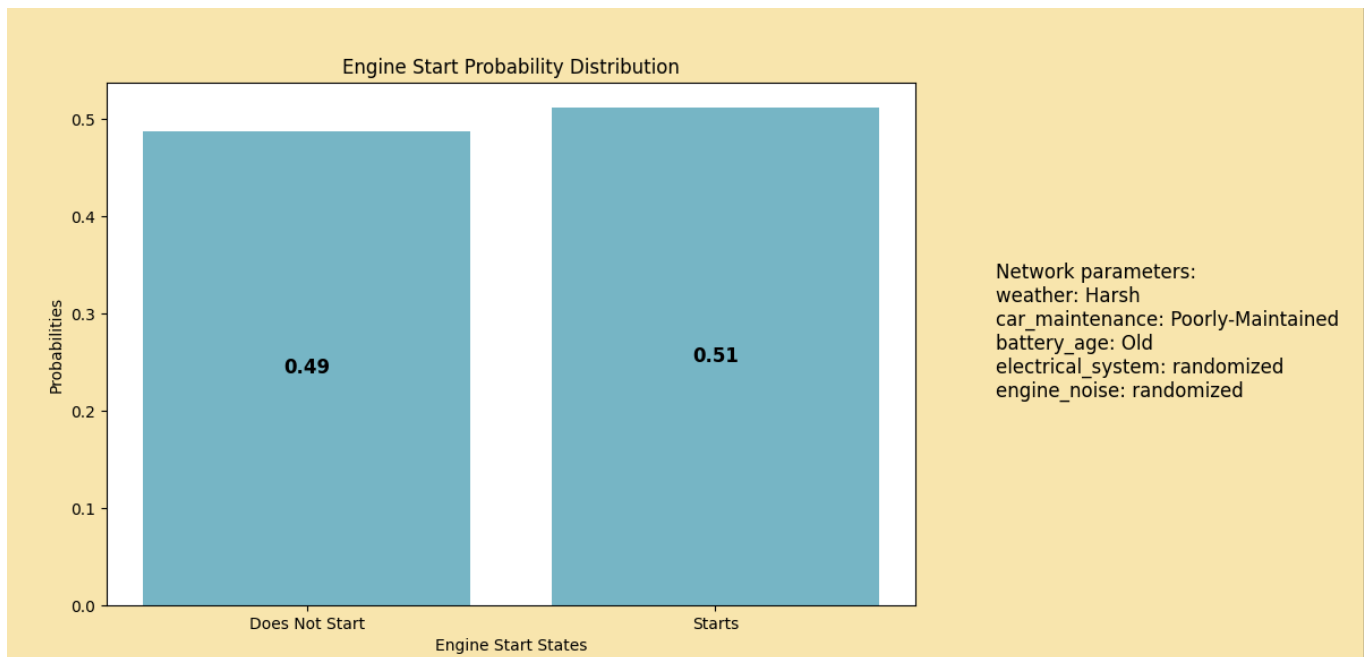
Rozpatrzmy trzy przykładowe sytuacje, ustalając wyłącznie prawdopodobieństwa w wierzchołkach sieci, które nie mają rodzica:

1. Pogoda sprzyjająca, nowy akumulator, aktualny serwis (optymistyczne ustawienie)

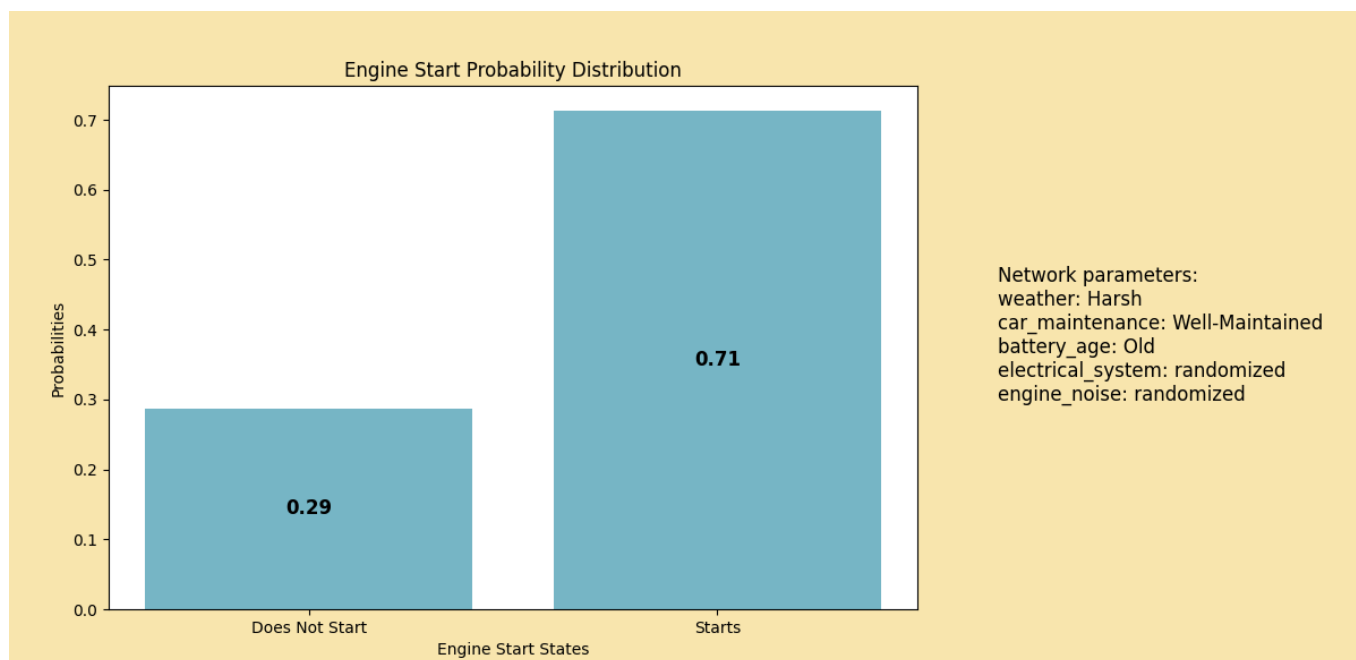


2. Pogoda niesprzyjająca, stary akumulator, brak

aktualnego serwisu (pesymistyczne ustawienie)



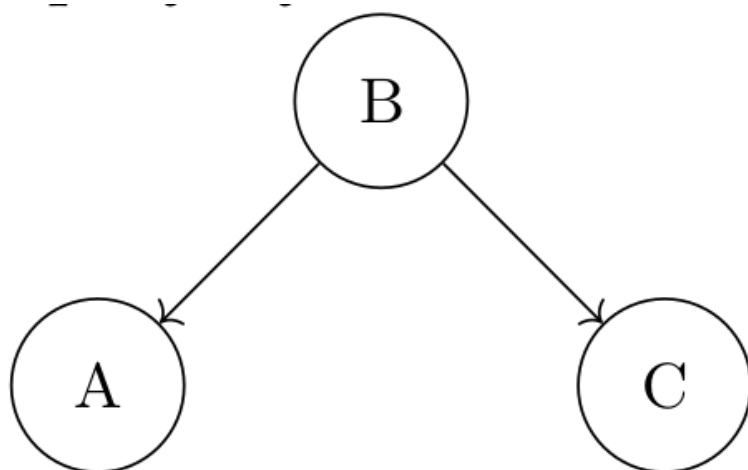
3. Niesprzyjająca pogoda, stary akumulator, serwis aktualny



Jak widzimy, ustawienie poszczególnych parametrów na „pozytywne” zwiększa prawdopodobieństwo na poprawne uruchomienie samochodu.

Warunkowa niezależność zmiennych w sieci

1. Dwie zmienne ze wspólną przyczyną



W takiej relacji w naszej sieci jest trójka serwis, działanie układu elektrycznego, dźwięk silnika. Rozpatrzmy, jaki rozkład prawdopodobieństwa ma dźwięk silnika, jeżeli określimy wyłącznie funkcjonowanie układu elektrycznego, a ich wspólna przyczyna pozostanie niewiadomą:

Układ elektryczny	P(dźwięk = normalny)	P(dźwięk nienormalny)
Funkcjonuje	77%	23%
Nie funkcjonuje	69%	31%

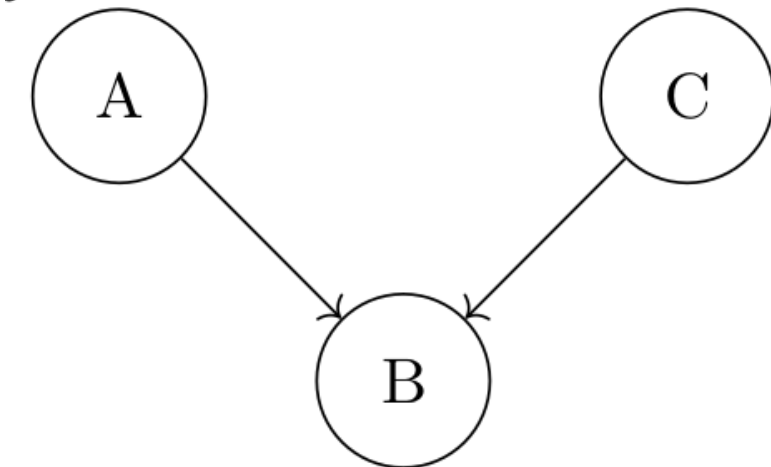
Widzimy, że dźwięk silnika jest zależny warunkowo od stanu układu elektrycznego. We wszystkich tych eksperymentach ustalamy jedynie interesujące nas zmienne, inne pozostawiamy do obliczenia. W tym wypadku sieć została uruchomiona 2 razy z różnymi wartościami dla układu elektrycznego.

Rozpatrzmy, co wydarzy się, gdy ustalimy zarówno stan układu, jak i stan serwisu. Przeanalizujemy wszystkie 4 kombinacje wartości tych dwóch zmiennych i sprawdzimy, jak zmienia się rozkład dla dźwięku.

Serwis	Układ elektryczny	P(dźwięk=normalny)	P(dźwięk=nienormalny)
Przeprowadzony	Funkcjonalny	85%	15%
Przeprowadzony	Niefunkcjonujący	85%	15%
Nieprzeprowadzony	Funkcjonujący	60%	40%
Nieprzeprowadzony	Niefunkcjonujący	60%	40%

Jak widzimy, po ustaleniu wartości serwisu, rozkład prawdopodobieństwa dźwięku zależny jest jedynie od serwisu, mimo że wcześniej zmieniał się w zależności od wartości stanu układu elektrycznego. Zaobserwowaliśmy więc, że zmienne te są w tej sytuacji warunkowo niezależne.

2. Dwie zmienne ze wspólnym skutkiem



Rozpatrzmy to na przykładzie trójki układ elektryczny, wiek baterii, uruchomienie silnika. Ustalmy najpierw wartości układu elektrycznego i sprawdźmy, czy zależy od nich rozkład wieku baterii:

Układ	P(bateria=nowa)	P(bateria=stara)
funkcjonalny	50%	50%
niefunkcjonalny	50%	50%

Wiek baterii nie jest zależny od stanu układu elektrycznego. Rozpatrzmy, co stanie się, gdy ustalimy wartość ich wspólnego skutku.

Układ	Uruchomienie silnika	P(bateria=nowa)	P(bateria=stara)
funkcjonalny	udane	55%	45%
funkcjonalny	nieudane	26%	73%
niefunkcjonalny	udane	56%	44%
niefunkcjonalny	nieudane	38%	62%

Jak widać, mamy cztery różne rozkłady stanu baterii, więc jest zależna nie tylko od uruchomienia silnika, ale też przez poznanie informacji o uruchomieniu, uwrażliwiła się na dane o układzie elektrycznym.

Dodatkowe przemyślenia

Warto na końcu powiedzieć, że struktura mojej sieci była mocno uproszczona, a prawdopodobieństwa najpewniej mogłyby być przydzielone sensowniej. Wydaje mi się jednak, że to nie realizm jest najważniejszy, a sensowność ogólnych prawidłowości oraz empiryczne zaobserwowanie reguł wnioskowania.