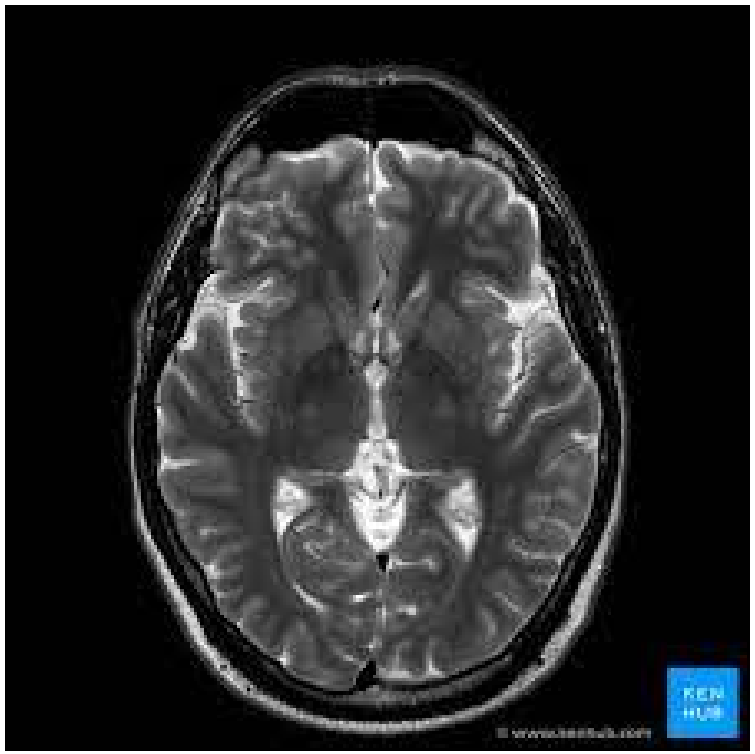
Conformal Prediction

Michał Kubica

16 kwietnia 2022

Motywacje

1. Sieć neuronowa
2. Diagnoza: Nowotwór
3. A może to jednak nie jest nowotwór?
4. Kolejne badania/nietrafiona diagnoza



Przykład

Chcemy ocenić toksyczność leku x_{n+1}. Mamy do dyspozycji:

1. Obserwacje historyczne $(x_1, x_2), \ldots, (x_n, y_n)$
   i x_i opisuje molekułę jako wektor numeryczny (predyktor)
   ii y_i - miara toksyczności. $\{\textit{toksyczny}, \textit{bezpieczny}\}$. Albo liczba przedziału $(0, 1)$
2. Model regresyjny lub klasyfikacyjny. NN, RF, SVM, etc.

Przykład c.d.

```
from sklearn.neighbors import KNeighborsClassifier
knn = KNeighborsClassifier(n_neighbors=5)
knn.fit(X_train, y_train)
print(knn.predict(X_test))
print(knn.predict_proba(X_test))
```

» [’bezpieczny’]
» [[0.8 0.2]]

Przykład c.d.

- Model twierdzi, że lek jest bezpieczny.
  > Czy jest tak na pewno?
- Model estymuje, że p-stwo, że lek jest bezpieczny wynosi 80%
  > Jak dobry jest ten estymator?
- Model regresyjny mówi, że lek ma poziom toksyczność 0.4
  > Jak blisko jesteśmy wartości prawdziwej?
- Czy można ufać takiemu modelowi?

Przykład c.d.

Odpowiedź: oczekujemy, że zachowanie w przyszłości będzie takie samo jak w przeszłości.

- Model ma 80% celności na zbiorze testowym.
  > Zakładamy, że model będzie miał 80% celności na produkcji.
- Model ma 0.9 AUC celności na zbiorze testowym.
  > Zakładamy, że model będzie miał 0.9 AUC na produkcji.
- Model ma RMSE na poziomie 0.2 na zbiorze testowym.
  > Zakładamy, że model będzie miał RMSE na poziomie 0.2 na produkcji.

ALE! Jak dobre są te estymatory? Czy mamy jakieś gwarancje/przedziały ufności? W szczególności, czy mamy ograniczenia górne/dolne na bezpieczeństwo leku x_{n+1}.

Conformal Prediction: Wstęp

- Przekształca klasyfikatory i regresory w predyktory przedziałowe (confidence predictors).
- Predykcje to raczej zbiory lub przedziały, a nie punkty.
  - > Klasyfikacja - np. $\hat{y} \in$ {guz, nowotwór, tętniak}
  - > Regresja - np. [0.3,0.43]
- Predykcje są ściśle związane z określonym poziomem ufności (confidence) - α

Conformal Prediction: Własności

- Otrzymujemy ograniczenia błędu osobno dla każdej obserwacji.
- Możemy tę metodę stosować dla jakiegokolwiek modelu uczenia maszynowego.
- Możemy tę metodę stosować dla dowolnego zbioru danych, w którym kolejność obserwacji nie ma znaczenia.
- Możemy tę metodę stosować do rozwiązań online i offline.

Conformal Prediction: Cele

- "Coverage"

$$\mathbb{P}[y_{n+1} \in T(X_{n+1})] \geq 1 - \alpha$$

- Mała moc zbioru T - mało elementów zbioru T
- Adaptacyjność - Moc T mała dla łatwych przykładów i duża dla trudnych przykładów.