W14. Psychoinformatyka w zastosowaniach klinicznych

Psychoinformatyka | Wykład 14

2026-11-04

Cele wykładu

- Odróżnij przesiew, diagnozę i prognozowanie.
- Zdefiniuj wynik oraz horyzont czasowy.
- Oceń kalibrację i przydatność modelu.

Obszary psychoinformatyki klinicznej

- Łączy samoopis, obserwację i biosygnały.
- Wymaga kontroli jakości danych.
- Wynik musi mieć znaczenie dla określonego celu.

Psychopatologia obliczeniowa

- Formalizuje hipotezy o mechanizmach zaburzeń.
- Parametry modeli mogą opisywać uczenie i decyzje.
- Nie stanowią samodzielnych rozpoznań.

Modele kategorialne i wymiarowe

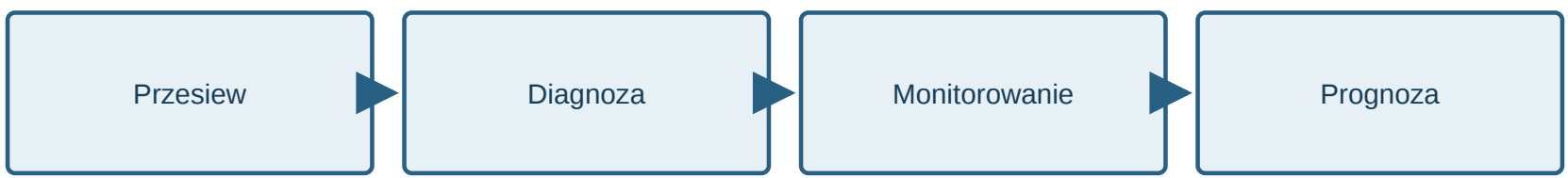
- Kategorie wspierają decyzje operacyjne.
- Wymiary opisują stopień nasilenia.
- Próg przesiewowy wymaga uzasadnienia.

RDoC i transdiagnostyka

- RDoC bada domeny funkcjonowania.
- Łączy różne poziomy pomiaru.
- Nie zastępuje systemu rozpoznań klinicznych.

Przesiew, diagnoza, monitorowanie, prognoza

- Przesiew wskazuje potrzebę dalszej oceny.
- Diagnoza wymaga kryteriów klinicznych.
- Prognoza dotyczy przyszłego wyniku.



Rysunek 1: Przesiew, diagnoza, monitorowanie, prognoza

Kwestionariusze PHQ-9 i GAD-7

- Skale samoopisowe opisują nasilenie zgłaszanych objawów.
- Wynik zależy od okresu odniesienia.
- Nie wystarcza do samodzielnej diagnozy.

Cyfrowe biomarkery i fenotypowanie

- Wskaźnik cyfrowy wymaga oceny analitycznej.
- Znaczenie kliniczne wymaga niezależnej walidacji.
- Korelacja nie przesądza o przyczynowości.

Wywiad, samoopis i dokumentacja

- Źródła danych mają różną trafność.
- Dokumentacja zależy od praktyki ośrodka.
- Zapisuj czas i pochodzenie etykiety.

Aktywność i mobilność

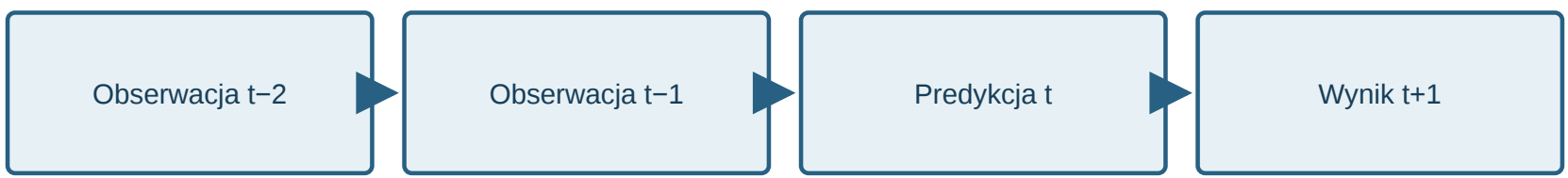
- Dane pasywne opisują zachowania, nie rozpoznania.
- Telefon i warunki pomiaru zmieniają wynik.
- Brak danych może być informacyjny.

Biosygnały i stan pacjenta

- HRV i EDA są niespecyficznymi wskaźnikami.
- Stan somatyczny wpływa na odczyt.
- Przenoszenie wyników wymaga walidacji.

Dane podłużne

- Jedna osoba dostarcza wiele zależnych obserwacji.
- Wynik w tygodniu t+1 jest przyszły.
- Predyktory muszą pochodzić sprzed prognozy.



Rysunek 2: Dane podłużne

Jakość etykiet klinicznych

- Rozpoznanie i wynik kwestionariusza to różne cele.
- Etykieta może być opóźniona lub niepewna.
- Dokumentuj sposób weryfikacji.

Braki danych i selekcja

- Brak danych może wiązać się z objawami.
- Rekrutacja zmienia reprezentatywność.
- Imputację dopasowuj na treningu.

Studium: wynik przyszłego tygodnia

- Prognozujemy fikcyjny podwyższony wynik przesiewu.
- Dane z tygodnia t poprzedzają t+1.
- Przykład obejmuje wyłącznie symulowane osoby.

5 minut przerwy

1. Czym przesiew różni się od diagnozy?
2. Dlaczego wynik skali nie jest rozpoznaniem?
3. Jak uniknąć przecieku danych przyszłych?
4. Co oznacza walidacja zewnętrzna?

Architektura wspomagania decyzji

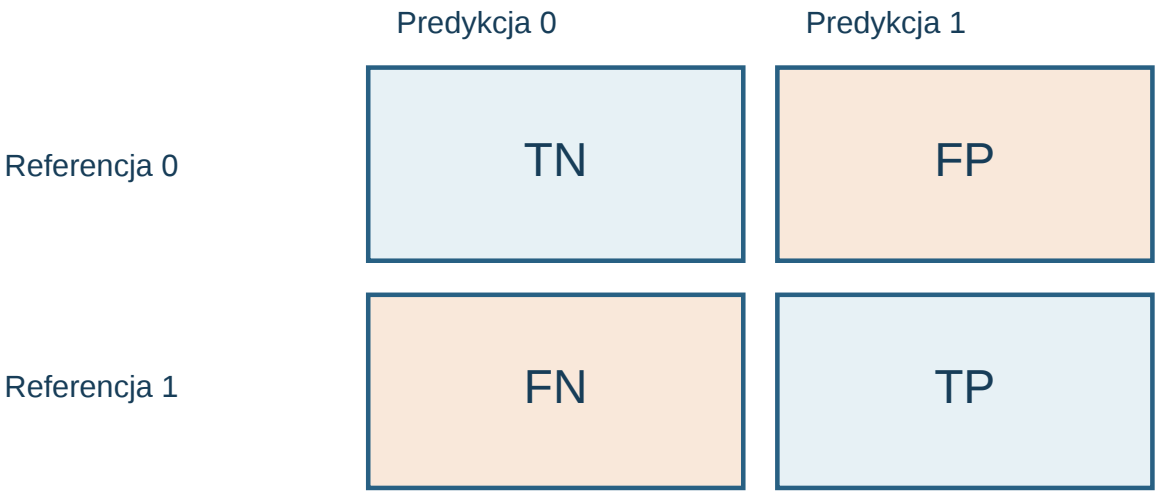
- Kontrola jakości poprzedza predykcję.
- Model dostarcza oszacowanie ryzyka.
- Specjalista interpretuje wynik.

Model ryzyka: populacja i wynik

- Określ populację i horyzont prognozy.
- Predyktory muszą być dostępne w punkcie decyzji.
- Utrwal kryterium wyniku przed analizą.

Czułość, swoistość, PPV i NPV

- Czułość dotyczy wykrywania przypadków dodatnich.
- Swoistość dotyczy poprawnych ujemnych.
- PPV i NPV zależą także od rozpowszechnienia.



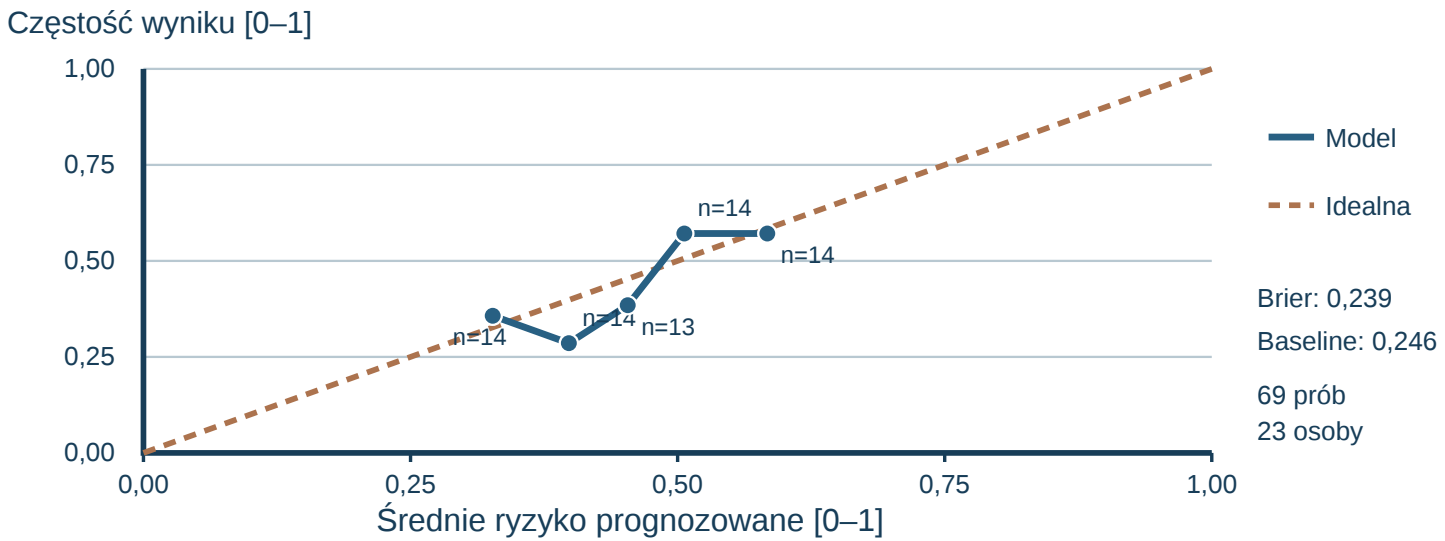
Rysunek 3: Czułość, swoistość, PPV i NPV

Progi i błędy

- Próg wpływa na FP i FN.
- Pominięcie i fałszywy alarm mają inne konsekwencje.
- Wybór wymaga jawnej funkcji kosztu.

Kalibracja ryzyka

- Dyskryminacja nie oznacza kalibracji.
- Brier mierzy błąd prawdopodobieństw.
- Krzywa kalibracji ocenia zgodność ryzyka i częstości.



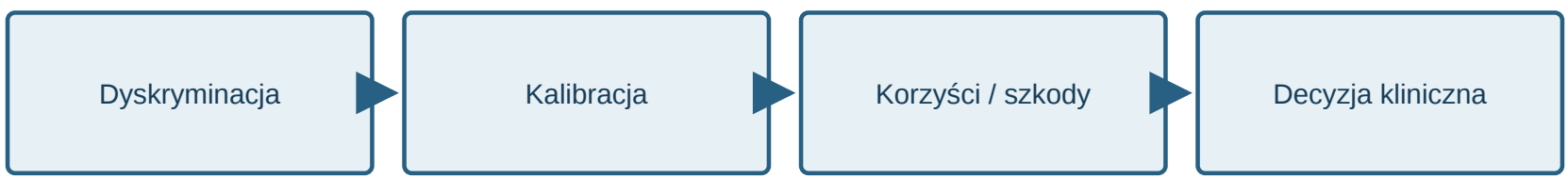
Rysunek 4: Kalibracja ryzyka

Walidacja zewnętrzna

- Sprawdź wynik na innej populacji lub ośrodku.
- Oceń zmianę rozpowszechnienia i urządzeń.
- Raportuj kalibrację i niepewność.

Przydatność kliniczna

- Decision curve analysis uwzględnia koszty błędów.
- Dyskryminacja nie oznacza korzyści netto.
- Wdrożenie wymaga oceny wpływu na decyzje.



Rysunek 5: Przydatność kliniczna

Python: syntetyczne dane podłużne

- Wielu uczestników ma cztery tygodnie obserwacji.
- Cechy z tygodnia t poprzedzają etykietę z tygodnia t+1.
- Etykieta wynika z fikcyjnego generatora.

```
import numpy as np
import pandas as pd
rng = np.random.default_rng(14)
osoba = np.repeat(np.arange(90), 4)
tydzien = np.tile(np.arange(4), 90)
cecha = rng.normal(size=90)[osoba]
sen = 7.2 - .35*cecha + rng.normal(0,.8,len(osoba))
aktywnosc = 6300 - 650*cecha + rng.normal(0,1500,len(osoba))
logit = -.5 + .8*cecha + .4*(7.2-sen) - .0001*(aktywnosc-6300)
wynik_t = rng.binomial(1, 1/(1+np.exp(-logit)))
dane = pd.DataFrame({"pid":osoba,"tydzien":tydzien,
                     "sen":sen,"aktywnosc":aktywnosc,"wynik_t":wynik_t})
dane["cel_t1"] = dane.groupby("pid")["wynik_t"].shift(-1)
dane["czas_celu"] = dane.groupby("pid")["tydzien"].shift(-1)
dane = dane.dropna(subset=["cel_t1"]).reset_index(drop=True)
assert len(dane) == 90*3
assert np.all(dane["czas_celu"] == dane["tydzien"]+1)
assert dane.groupby("pid")["tydzien"].max().eq(2).all()
pid = dane["pid"].to_numpy()
y = dane["cel_t1"].astype(int).to_numpy()
X = dane[["sen","aktywnosc","tydzien"]]
```

Python: podział grupowy

- Podziel dane według identyfikatora osoby.
- Imputacja i skalowanie są częścią pipeline’u.
- Test jest rozłączny osobowo: walidacja wewnętrzna.

```
from sklearn.model_selection import GroupShuffleSplit
from sklearn.pipeline import make_pipeline
from sklearn.impute import SimpleImputer
from sklearn.preprocessing import StandardScaler
from sklearn.linear_model import LogisticRegression
split = GroupShuffleSplit(n_splits=1,test_size=.25,random_state=17)
tr,te = next(split.split(X,y,groups=pid))
assert set(pid[tr]).isdisjoint(set(pid[te]))
model = make_pipeline(SimpleImputer(),StandardScaler(),LogisticRegression())
model.fit(X.iloc[tr],y[tr]); p = model.predict_proba(X.iloc[te])[:,1]
```

Python: ROC-AUC i Brier

- Oblicz dyskryminację na teście.
- Porównaj Brier modelu z baseline opartym na treningu.
- Raportuj liczebność uczestników.

```
from sklearn.metrics import roc_auc_score,brier_score_loss
p_base=np.full(len(te),y[tr].mean(),dtype=float)
auc=roc_auc_score(y[te],p)
brier=brier_score_loss(y[te],p)
brier_base=brier_score_loss(y[te],p_base)
print("Osoby testowe:",len(np.unique(pid[te])))
print("ROC-AUC:",round(auc,3))
print("Brier model / baseline:",round(brier,3),round(brier_base,3))
print("Częstość train/test:",round(y[tr].mean(),3),round(y[te].mean(),3))
rng_ci=np.random.default_rng(140)
osoby_test=np.unique(pid[te]);ci_auc=[]
for _ in range(300):
    osoby_bs=rng_ci.choice(osoby_test,len(osoby_test),replace=True)
    ix=np.concatenate([np.flatnonzero(pid[te]==o) for o in osoby_bs])
    if len(np.unique(y[te][ix]))==2:
        ci_auc.append(roc_auc_score(y[te][ix],p[ix]))
print("95% CI ROC-AUC, bootstrap osób:",
      np.round(np.percentile(ci_auc,[2.5,97.5]),3))
```

Python: progi i błędy

- Z góry ustal trzy progi.
- Porównaj TN, FP, FN i TP.
- Wizualizuj czułość i empiryczną krzywą kalibracji.

```
from sklearn.metrics import confusion_matrix
from sklearn.calibration import calibration_curve
import matplotlib.pyplot as plt
progi=[.35,.50,.65];czul=[]
for prog in progi:
    tn,fp,fn,tp=confusion_matrix(y[te],p>=prog,labels=[0,1]).ravel()
    czul.append(tp/(tp+fn));print(prog,tn,fp,fn,tp)
czestosc,ryzyko=calibration_curve(y[te],p,n_bins=5,strategy="quantile")
fig,ax=plt.subplots(1,2,figsize=(10,3.5))
ax[0].plot(progi,czul,marker="o")
ax[0].set(xlabel="Próg",ylabel="Czułość [0–1]",ylim=(0,1))
ax[1].plot([0,1],[0,1],"--",label="Idealna")
ax[1].plot(ryzyko,czestosc,marker="o",label="Model")
ax[1].set(xlabel="Średnie ryzyko",ylabel="Częstość [0–1]",
          xlim=(0,1),ylim=(0,1))
ax[1].legend()
for panel in ax:panel.grid(alpha=.3)
plt.tight_layout();plt.show()
```

TRIPOD+AI i PROBAST+AI

- TRIPOD+AI opisuje raportowanie modeli klinicznych.
- PROBAST+AI wspiera ocenę stronniczości.
- Raportuj braki, dobór osób i walidację.

Bezpieczeństwo i ograniczenia

- Przewiduj błędy, eskalację i kontrolę człowieka.
- Monitoruj działanie w kolejnych populacjach.
- Skuteczność populacyjna nie gwarantuje trafności jednostkowej.

Literatura i dokumentacja

- Collins i in. (2024) – TRIPOD+AI.
- Moons i in. (2025) – PROBAST+AI.
- Dokumentacja scikit-learn opisuje metryki.

TRIPOD+AI (2024) · PROBAST+AI (2025).

Wnioski i samodzielna praca

- Przesiew nie jest diagnozą.
- Sprawdzaj dyskryminację, kalibrację i bezpieczeństwo.
- Uruchom pi_w14.ipynb i porównaj progi.