

Beleidsrapport: Voorspelling van Ziekenhuisopname op Basis van Gezondheidskenmerken

Inleiding

In dit project is onderzocht in hoeverre we het risico op ziekenhuisopname, benaderd via diabetesstatus, kunnen voorspellen op basis van gezondheids- en leefstijlkenmerken. Het doel was om met behulp van machine learning inzichten te verkrijgen die bruikbaar zijn voor preventief gezondheidsbeleid.

De dataset bevatte meer dan 250.000 waarnemingen uit de Amerikaanse BRFSS 2015 survey. Als proxy voor ziekenhuisopname is gekozen voor de variabele Diabetes_binary, die aangeeft of iemand ooit de diagnose diabetes type 2 heeft gekregen.

Samenvatting van modelresultaten

Twee modellen zijn getraind en geëvalueerd:

Model	Precision	Recall	F1-score	ROC AUC
Logistic Regression	0.75	0.75	0.75	0.82
Random Forest	0.74	0.74	0.74	0.81

Interpretatie:

Beide modellen presteren vergelijkbaar en tonen een goed onderscheidend vermogen (ROC AUC > 0.80). De modellen zijn dus geschikt om risicogroepen te identificeren op basis van gezondheidskenmerken.

Belangrijkste voorspellende kenmerken

De Random Forest identificeerde de volgende top 10 factoren voor het voorspellen van verhoogd risico:

- BMI
- Leeftijd
- Gepercipieerde gezondheidstoestand

4. Inkomen
5. Hoge bloeddruk
6. Gepercipieerde fysieke gezondheid
7. Opleidingsniveau
8. Gepercipieerde mentale gezondheid
9. Hoge cholesterol
10. Rookgedrag

Deze factoren zijn een mix van objectieve (zoals BMI en bloeddruk) en subjectieve gezondheidsindicatoren, wat erop wijst dat zelfperceptie van gezondheid ook een belangrijke rol speelt.

Beleidsaanbevelingen

Op basis van deze inzichten kunnen de volgende beleidsrichtingen worden overwogen:

1. Preventie gericht op overgewicht en obesitas

- BMI is de sterkste voorspeller.
- ► Investeer in laagdrempelige voedings- en beweeginterventies (schoolprogramma's, buurtsportcoaches, voedingsvoorlichting).

2. Screening en monitoring bij risicogroepen

- Hoge bloeddruk en cholesterol zijn relevante signalen.
- ► Stimuleer jaarlijkse gezondheidschecks bij mensen met overgewicht of lage SES.

3. Versterken van gezondheidsvaardigheden

- Gepercipieerde gezondheidstoestand, mentale gezondheid en opleidingsniveau zijn sterk gelinkt aan risico.
- ► Zet in op gezondheidseducatie, taalondersteuning, en mentale gezondheidszorg op wijkniveau.

4. Targeting op sociaaleconomische status (SES)

- Inkomen en opleidingsniveau komen sterk naar voren.
- ➤ Ontwikkel specifieke campagnes voor lage-inkomenswijken of werklozen.

5. Rookstop-programma's uitbreiden

- Roken blijft voorspellend, maar is beïnvloedbaar.
- ➤ Maak stoppen-met-roken-coaching (digitaal of live) gratis voor risicogroepen.

Conclusie

Machine learning laat zien dat met relatief eenvoudige data al zinvolle voorspellingen kunnen worden gedaan over gezondheidsrisico's. Dit soort modellen kunnen beleidsmakers ondersteunen in het gericht inzetten van preventiemiddelen bij de juiste doelgroepen, en zo bijdragen aan lagere ziekenhuisopname en betere volksgezondheid.