

Srdeční onemocnění¹

Teoretické ukotvení

Wenger, Speroff, Packard, & Shapiro (1984) se věnovali různým faktorům ovlivňující výskyt, vývoj a progresi srdečních chorob, včetně věku, pohlaví, krevního tlaku a přítomnosti anginy. Studie ukázala, že tyto faktory, spolu s dalšími sociodemografickými a psychologickými faktory, mohou přispět k vývoji a progresi srdečních chorob. Zároveň zdůrazňují potřebu multidisciplinárního přístupu k prevenci a léčbě této nemoci.

V posledních letech vznikají studie, které se pokouší predikovat výskyt srdečního onemocnění pomocí speciálních algoritmů. Aha a Kibler (1991) se zabývali využitím algoritmů učení k předpovídání přítomnosti srdečních onemocnění. Studie ukázala, že algoritmy učení na základě instancí dosahují lepších výsledků než jiné algoritmy strojového učení při přesném předpovídání přítomnosti srdečních onemocnění u pacientů. Detrano a spol. (1989) popisuje použití pravděpodobnostního algoritmu k diagnostikování koronární arteriální choroby (CAD). Studie ukázala, že algoritmus je účinný při identifikaci přítomnosti a závažnosti CAD.

Binomiální logistická regrese

Binomiální logistická regrese je statistický model používaný k zkoumání vztahu mezi jedním nebo více proměnnými a binárními odpověďmi. Model odhaduje pravděpodobnost úspěchu nebo selhání binární odpovědní proměnné na základě hodnot nezávislých proměnných (Hosmer, Lemeshow a Sturdivant, 2013). V této práci autorka pomocí binomiální logistické regrese ověřuje možnost předpovídat výskyt diagnózy srdečního onemocnění (závislá proměnná) na základě více nezávisle proměnných. Konkrétně byly do modelu zařazeny následující proměnné:

Závisle proměnná = (ne)přítomnost diagnózy srdečního onemocnění u pacienta

Nezávislé proměnné:

- **Věk.**
- **Pohlaví** (biologické, muž/žena).
- **Typ bolesti na hrudi (angina):**
 - 0 = typická angina
 - 1 = atypická angina
 - 2 = neanginózní
 - 3 = asymptomatická
- **Klidový krevní tlak.**
- **Sérový cholesterol v mg/dl.**
- **Hladina glukózy v krvi > 120 mg/dl.**
- **Výsledky EKG v klidu:**
 - 0 = normální
 - 1 = s pravděpodobnou nebo možnou abnormalitou
 - 2 = s definitivními důkazy o abnormalitě
- **Maximální dosažená srdeční frekvence.**
- **Výskyt bolesti na hrudi způsobený cvičením.**
- **Oldpeak** = stupeň deprese v segmentu ST elektrokardiogramu po cvičení X klidovým ST segmentem.

¹ Tato zpráva vznikla v rámci předmětu Vícerozměrné statistické metody (K. psychologie FF UPOL, 2022/23) Data a další informace o této zprávě jsou dostupné na adrese: <https://dostal.vyzkum-psychologie.cz/stat4?i=6>

- **Počet hlavních cév (0-3) obarvených fluoroskopií.**
- **Klasifikace abnormalit prokrvení myokardu určených Thalliovou scintigrafií:**
 - 0 = normální prokrvení
 - 1 = fixní defekt
 - 2 = reverzibilní defekt

Na datovém souboru o 304 participantech byla provedena binomiální logistická regrese týkající se proměnných vypsanych výše. V tabulce níže (Tab. 1) jsou zaznamenány konkrétní výsledky. Pro účely testové statistiky byla zvolena Waldova statistika, dále je možné v tabulce pozorovat hodnoty regresivních koeficientů a příslušné p hodnoty.

Tab. 1: Výsledky binomiální logistické regrese

Model	Regresní koeficient	Testová statistika	p hodnota
Věk	1,00	0,00	0,98
Pohlaví	4,55	8,45	0,00
Angina typ 0	0,13	9,60	0,00
Angina typ 1	0,36	1,84	0,18
Angina typ 2	0,93	0,01	0,91
Angina typ 3	5,01	19,63	0,00
Klidový krevní tlak.	0,98	2,55	0,11
Sérový cholesterol	1,00	1,24	0,27
Hladina glukózy v krvi > 120 mg/dl.	0,84	0,10	0,76
Výsledky EKG - 0	1,32	0,01	0,90
Výsledky EKG - 1	2,33	0,14	0,71
Výsledky EKG - 2	0,97	2,38	0,30
Maximální dosažená srdeční frekvence.	1,02	2,55	0,11
Výskyt bolesti na hrudi způsobený cvičením.	2,14	3, 21	0,07
Oldpeak	0,61	4,70	0,03
Počet hl. cév obarvených fluoroskopií.	0,43	16,63	0,00
Thaliova scintigrafie - 0	0,62	0,04	0,84
Thaliova scintigrafie - 1	3,82	3,13	0,08
Thaliova scintigrafie - 2	3,98	11,32	0,00

* šedě jsou zvýrazněny signifikantní výsledky

Signifikantní výsledek je prokázán u 6 proměnných. Jedná se o vliv pohlaví, přítomnost typické anginy, přítomnost asymptotické anginy, oldpeak, počet hl. cév obarvených fluoroskopií a reverzibilní defekt prokrvení myokardu určených Thalliovou scintigrafií. Tyto proměnné mají velmi malou p hodnotu, a zároveň Waldova statistika dosahuje hodnot vysokých. Na základě těchto výsledků můžeme právě o těchto proměnných uvažovat jako o významných v souvislosti s detekcí srdečního onemocnění, a tak i významných pro další rozvoj detekce srdečních onemocnění.

Ukazatelé kvality modelu

Pro zjištění kvality modelu jsou využity ukazatele Nagelkerke R^2 a Cox-Snell R^2 . Hodnota Cox-Snell R^2 je rovna 0,51, tedy vysvětluje přibližně 51% a hodnota Nagelkerke R^2 je rovna 0,68, tedy vysvětluje přibližně 68%. Celkově tedy tyto ukazatele naznačují, že binomiální logistický regresní model má střední až dobrou přizpůsobivost k datům a může být použit k vytvoření poměrně přesných predikcí.

Literatura

Aha, D. W. a Kibler, D. (1991). Instance-based prediction of heart-disease presence with the Cleveland database [Abstrakt]. Získáno z <https://www.ncbi.nlm.nih.gov/pubmed/1769732>

Detrano, R., Janosi, A., Steinbrunn, W., Pfisterer, M., Schmid, J., Sandhu, S., Guppy, K., Lee, S. a Froelicher, V. (1989). International application of a new probability algorithm for the diagnosis of coronary artery disease. *American Journal of Cardiology*, 64, 304-310.

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. John Wiley & Sons.

Wenger, N. K., Speroff, L., Packard, B. a Shapiro, J. (1984). Koronární srdeční choroba: sociodemografický, psychologický a lékařský profil. *Circulation*, 69(2), 237-244. <https://doi.org/10.1161/01.CIR.69.2.237>