

Predispozice k vývoji srdeční choroby¹

Teoretické ukotvení

Kardiovaskulární onemocnění je souhrnný medicínský termín označující choroby postihující srdce a krevní cévy. Tento soubor onemocnění představuje zhruba 32 % všech úmrtí na světě, což z něho dělá hlavní příčinu úmrtnosti vůbec. Zajímavostí je, že cca 75 % úmrtí způsobené kardiovaskulárními onemocněními se dějí v rozvojových zemích. Vývoj nemoci probíhá za situace, ve které cévy zásobující srdce krví a kyslíkem, začnou tuhnout a zužovat se kvůli nahromadění tukových látek (Pravesh, 2022).

Mezi významné prediktory vzniku kardiovaskulárního onemocnění dle Abd El-Wahaba (2021) patří stáří, fyzická neaktivita, špatné stravovací návyky spojené s nadměrnou konzumací vysoce zpracovaných potravin, nadměrné užívání alkoholu a kouření mimo jiné.

Jelikož tato skupina nemocí stojí na první příčce, co se týče příčin úmrtí, je kladen velký důraz na výzkum prediktorů a chování způsobující kardiovaskulární potíže. Novodobým trendem je využívat takzvaný *machine learning* (ML) čili strojové učení (v zásadě o využití umělé inteligence). Je to metoda umožňující strojové technice „se učit“ za účelem zvýšení přesnosti predikce rozvoji kardiovaskulárních potíží. (Alaa et al., 2019)

Způsob zpracování dat

Data budeme zpracovávat binomální logistickou regresí, která se používá v situacích, ve kterých zkoumáme vztah mezi alternativní závislou proměnnou a jednou či více nezávislými proměnnými. Závislá proměnná nabývá tedy pouze hodnot 1 a 0, v tomto případě představuje **závislá proměnná**, zda u pacienta byla diagnostikována srdeční choroba či nikoli.: **1 = byla diagnostikována srdeční choroba, 0 = nebyla diagnostikována srdeční choroba** (Hosmer & Lemeshow, 2000).

Nezávislé proměnné použité v modelu jsou následující:

Věk

Pohlaví (biologické)

¹ Data a další informace o této zprávě jsou dostupné na adrese <https://dostal.vyzkum-psychologie.cz/stat4?i=141>.

Bolest hrudníku prožívaná pacientem

Krevní tlak

Úroveň cholesterolu

Výsledek testu krevního cukru nalačno větší než 120 mg/dl

Výsledek EKG

Maximální dosažená tepová frekvence během cvičení

Angína projevená během zátěžového testování

Množství cév viditelných na fluoroskopii (0-3)

Výsledky zátěžové testu Thallium (test zjišťující, jak dobře krev proudí do srdce, pokud člověk cvičí nebo je v klidu)

Na datovém souboru čítajícím 270 respondentů byla provedena binomální logistická regrese s použitím proměnných podle výše uvedené deskripce. Tabulka č. 1 zobrazuje výsledky této analýzy, ve které byla použita Waldova statistika a rovněž jsou zahrnuty hodnoty regresivních koeficientů a p hodnoty.

Model	Regresní koeficient	Testová statistika	p hodnota
Věk	0,98	0,68	0,410
Pohlaví	4,55	8,58	0,003
Bolest hrudníku	1,92	9,57	0,002
Krevní tlak	1,03	6,43	0,011
Cholesterol	1,01	3,37	0,066
Krevní cukr	0,49	1,72	0,190
EKG	1,42	3,28	0,070
Tep. frekvence	0,97	8,30	0,004
Angína	2,62	5,16	0,023
Fluoroskopie	3,21	20,45	>0,001
Thallium test	1,46	13,60	>0,001

* šedou barvou jsou zvýrazněny signifikantní výsledky

Z výpočtů jsou patrné signifikantní výsledky u 7 proměnných, konkrétně jde o pohlaví, bolest hrudníku, krevní tlak, tepovou frekvenci, angínu, fluoroskopii a Thallium test. Všechny

tyto proměnné mají p hodnotu nižší než 0,05 a rovněž Waldova statistika dosahuje vyšších hodnot. Na základě těchto dat můžeme tyto signifikantní proměnné dát do souvislosti s rozvojem srdečních chorob.

Ukazatelé kvality modelu

Jako ukazatele kvality modelu byly vybrány ukazatele Cox-Snell R^2 a Nagelkerke R^2 . Výsledná hodnota Cox-Snell R^2 se rovná 0,49, vysvětluje tedy 49 %, Nagelkerke R^2 nabývá hodnoty 0,66, vysvětluje tak 66 %. Tato čísla se dají interpretovat tak, že prezentovaný regresní model má střední až vysokou přizpůsobivost k datům čili může být použit jako poměrně přesný prediktor.

Zdroje:

1. Abd El-Wahab, E. W. (2021). Predicting coronary heart disease using risk assessment charts and risk factor categories. *Journal of Public Health: From Theory to Practice*, 29(5), 1037-1045. <https://doi.org/10.1007/s10389-020-01224-z>
2. Alaa, A. M., van der Schaar, M., Bolton, T., Angelantonio, E. D., & Rudd, J. H. F. (2019). Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. *PLoS ONE*, 14(5). <https://doi.org/10.1371/journal.pone.0213653>
3. Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* / David W. Hosmer, Stanley Lemeshow.
4. Pravesh, R. (2022). Comparative Performance Analysis of Random Forests against AutoPrognosis for predicting Coronary Heart Disease Risk and Metabolic Syndrome: A Retrospective Cohort Study. *MATEC Web of Conferences*, 370, 07005-7005. <https://doi.org/10.1051/mateconf/202237007005>