

Autor: Bc. Matej Čajka

cajka10@stud.uniza.sk

Vedúci: doc. Ing. Ján Boháčik, PhD.

jan.bohacik@fri.uniza.sk

Školiteľ: Ing. Ivan Škula

skula4@stud.uniza.sk

DATAMINING PRI DETEKCII RAKOVINY PRSNÍKA



Žilinská univerzita v Žiline, Fakulta riadenia a informatiky

PROBLEMATIKA

DIAGNOSTIKA

V našej práci sme sa zamerali na najrozšírenejší typ rakoviny. Rakovinou prsníka trpí na Slovensku až **12% dospelaj populácie žien**, čo činí každú ôsmu ženu v našej krajine. Ročne pribudne až **3300 nových prípadov**. [1] Je to vážny problém a môžeme to považovať za **civilizačné ochorenie**.

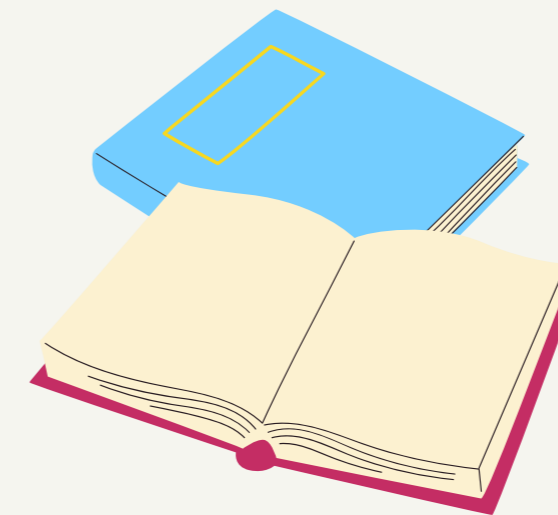
[1] Ondrušová, M., Pšenková, M., Suchanský, M.: Vybrané epidemiologické aspekty karcinómu prsníka. Bratislava, Pharm-In 2016



NÁSTROJE

JAVA/WEKA/PYTHON/SCI-KIT/POSTGRESQL

Pre vypracovanie projektu bol použitý programovací jazyk **Java**, platforma **JavaFX** použitá pre vývoj desktopovej aplikácie. Databázový systém **PostgreSQL** pre uchovávanie dát. Knižnica **WEKA** pre implementáciu algoritmov strojového učenia. **Python** a knižnice, ktoré taktiež slúžili na implementovanie strojového učenia a analýzu algoritmov strojového učenia.



VYUŽITIE V PRAXI

UŽIVATEĽSKÁ APLIKÁCIA

Vytvorili sme demo **desktopovú aplikáciu**, ktorá slúži priamo pre lekárov pre udržiavanie si záznamov o pacientoch a priame využitie modelov strojového učenia pri **predikcii zhubnosti nádoru**. Cieľom aplikácie je uľahčiť prácu lekárom, urýchliť a zefektívniť proces diagnostiky zhubnosti nádoru.

MEDICÍNA A STROJOVÉ UČENIE

DATAMINING

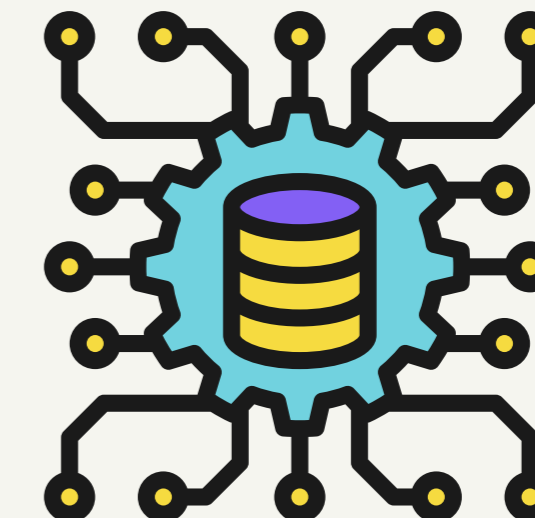
V zdravotníctve sa zbiera a spracováva **obrovské množstvo dát**, čo je predpoklad na efektívne využitie strojového učenia. Algoritmy strojového učenia sú v zdravotníctve využiteľné pri predikcii chorôb, symptómov, pre podporu rozhodovania a či už pri samotnej liečbe pacienta. Pri správnom použití je ideálny nástroj pre skvalitnenie zdravotníctva, ako na Slovensku, tak aj v celom svete.



IMPLEMENTÁCIA

RIEŠENIE

Pri práci sme využili voľne dostupnú knižnicu **WEKA**, na implementovanie algoritmov strojového učenia. Po analýze sme zvolili 2 algoritmy a natrénovali sme modely **J48** a **Multilayer Perceptron**, ktoré dosahovali **presnosť presahujúcu 90%** na náhodnej testovacej množine dát, pre nasledovné použitie klasifikácie priamo v aplikácii.



"The world is one Big Data problem" Dr. Andrew McAfee