

# VYUŽITIE PRVKOV UMELEJ INTELIGENCIE PRI KLASIFIKÁCIÍ EMOCIONÁLNEHO STAVU NA ZÁKLADE BEHAVIORÁLNEJ CHARAKTERISTIKY POUŽÍVATEĽA



UNIVERZITA  
KONŠTANTÍNA  
FILOZOFA  
V NITRE

KATEDRA INFORMATIKY  
FPVaI UKF v Nitre

Autor: Mgr. Dávid Držík; Školiteľ: doc. PaedDr. Martin Magdin, Ph.D.

## Cieľ práce

Hlavným cieľom diplomovej práce je využiť prvky umelej inteligencie pri klasifikácii emocionálneho stavu používateľa na dátach tvorených charakteristikami správania sa, získaných z vlastnej desktopovej a mobilnej aplikácie. Pri klasifikácii sme využívali vlastný dataset pozostávajúci aj z hodnotenia vlastného emocionálneho stavu používateľov na dimenziách valencie a vzrušenia v Russelovom kruhovom modeli emócií.

## Tvorba experimentov

V aplikačnej časti práce sme vytvorili viacero experimentov, so zameraním na prácu s klávesnicou, myšou a dotykovým mobilným zariadením. V aplikácii **Emotnizer desktop** (Java 8) sme implementovali 4 experimenty. V prvých dvoch bolo cieľom prepísať vopred definovaný text pomocou klávesnice počítača a v posledných dvoch bolo cieľom postláčať objekty pomocou myši počítača k splneniu úlohy daných experimentov.

V aplikácii **Emotnizer mobile** (Android Studio), sme implementovali 2 experimenty. V prvom išlo opäť o prepis preddefinovaného textu, avšak pomocou našej vlastnej softvérovej klávesnice. A v druhom experimente bolo cieľom posunúť štvorce hlavolamu "8 puzzle problem" pomocou prsta tak, aby boli usporiadané vzostupne. Na absolvovanie každého experimentu mali používatelia vymedzený čas.

Povinnosťou používateľov bolo pred aj po absolvovaní každého experimentu **ohodnotiť svoj emocionálny stav** na dimenziách valencie a vzrušenia v Russelovom kruhovom modeli emócií, na škále od -1 po 1.

Na konci každého experimentu sa z klikaní, stláčaní, držaní myši, klávesnice aj prsta na obrazovke **vypočítali charakteristiky**, ako napríklad frekvencia klikania, doba držania klávesu či pauza medzi klávesmi, a zo všetkých charakteristík sme analyzovali základnú štatistiku týchto charakteristík.

## Zber, analýza a predspracovanie dát

V období august až október roka 2021 sa našich experimentov zúčastnilo 51 osôb. Požadovali sme od nich, aby sa najskôr registrovali na našej webovej stránke. Potom sa pomocou emailu a hesla prihlásili do aplikácií a robili experimenty. Dáta získané z aplikácií zo všetkých šiestich experimentov sa posielali na server a cez PHP skripty sa vkladali dáta do MySQL databázy. Vo výsledku sme získali až **6 dátových súborov**.

Fázu analýzy, predspracovania, ale aj samotné nasadenie klasifikačných algoritmov strojového učenia sa sme robili v programovacom jazyku Python.



Následne sme všetky vstupné dáta zoškálovali a hodnotenia emócií po absolvovaní experimentov sme zmenili z podoby reálnych čísel do troch kategórií pre obe cieľové premenné. Premenná **valencie** obsahuje 3 stavy (negatívna, neutrálna a pozitívna) a premenná **vzrušenia** nadobúda tiež 3 stavy (nízke, stredné, vysoké). Všetky dátové súbory sme rozdelili na tréningové a testovacie množiny a pomocou krížovej validácie sme overili presnosť klasifikácie našich modelov.

## Modely strojového učenia

Aplikovali sme 6 algoritmov strojového učenia (**K-Najbližších susedov**, **Logistická regresia**, **Naivný Bayesov klasifikátor**, **Neurónová sieť**, **Random Forest** a **Support Vector Machine**) na všetkých šiestich dátových súboroch, čím nám vzniklo až 36 modelov pre obe cieľové premenné.

Pri vyhodnotení presnosti klasifikácie algoritmov strojového učenia sme vyhodnocovali presnosť klasifikácie modelu pomocou **confusion matrix** a metrík **accuracy** a **F1 score**.

Spomedzi všetkých modelov, najlepšiu klasifikačnú presnosť sme dosahovali modelom neurónovej siete klasifikovaním **vzrušenia** datasetom 2 na úrovni **87,88%** a klasifikovaním **valencie** datasetom 1 na úrovni **85,61%**.

**Sekvenčný model** neurónovej siete knižnice Keras pozostával zo vstupnej, z troch skrytých a z výstupnej vrstvy. V každej vrstve sme definovali počet neurónov a aktivačnú funkciu. Vytvorený model sme skompilovali, natrénovali a otestovali. Nasledujúce grafy predstavujú presnosť a chybovosť modelu počas jednotlivých iterácií.

