Predikcia budúcich nákladov na pacienta

Marián Kravec školiteľ: MSc. František Dráček

Predikcia budúcich nákladov na pacienta

Diel práce

Data

Embedo

ýkony

Lieky

Predikcia náklado

Proces predikcie

Výpočet počtu budúcio

záznamov Predikcia budúcich

záznamov Predikcia cenovej kategóri

Predikcia cenovej kategóri záznamu

testovanie a

validácia

Trénovanie a testovanie Validácia

Obsah Cieľ práce

Dáta

Embedding

Výkony Diagnózy

Predikcia nákladov

Proces predikcie Výpočet počtu budúcich záznamov Predikcia budúcich záznamov Predikcia cenovej kategórie záznamu

Trénovanie, testovanie a validácia

Trénovanie a testovanie Validácia

Ďalšie kroky

Predikcia budúcich nákladov na pacienta

Cieľ práce

Dáta

Embedo

/ýkony

Diagnózy Lieky

Predikcia nákladov

Proces predikcie

Výpočet počtu budúcie
záznamov

záznamov Predikcia budúcich

Predikcia cenovej kategório záznamu

testovanie a validácia

Trénovanie a testovanie

Cieľ práce

Predikovanie budúcej nákladovej skupiny do ktorej bude pacient v budúcnosti patriť na základe jeho minulosti a súčasného zdravotného stavu

Predikcia budúcich nákladov na pacienta

Cieľ práce



Cieľ práce

Dáta

Embedding

Výkony Diagnózy

Predikcia nákladov

Proces predikcie Výpočet počtu budúcich záznamov Predikcia budúcich záznamov Predikcia cenovej kategórie záznamu

Trénovanie, testovanie a validácia

Trénovanie a testovanie Validácia

Ďalšie kroky

Predikcia budúcich nákladov na pacienta

Cieľ práce

Dáta

Embodo

/ýkony

Lieky

Predikcia nákladov

Proces predikcie Výpočet počtu budúcich záznamov

Predikcia budúcich záznamov Predikcia cenovej kategór

redikcia cenovej katego iznamu

testovanie a validácia

Trénovanie a testovanie

Dáta

Predikcia budúcich nákladov na pacienta

Dáta

Zdroj dát: NCZI

Typ dát: poistné dáta

- Jeden záznam: informácia o jednom výkone/lieku konkrétneho pacienta
- Datasety:
 - Dataset ambulantnej zdravotnej starostlivosti
 - Dataset ústavnej zdravotnej starostlivosti
 - Dataset predpísaných liekov
- V súčastnosti využívaná malá anonymizovaná vzorka



Najdôležitejšie údaje pre časový rad

- Čas od prvého výkonu
- Vek pacienta
- Medicínsky výkon
- Diagnóza
- Liek
- Cena výkonu/lieku

Predikcia budúcich nákladov na pacienta

Dáta



Cieľ práce

Dáta

Embedding

Výkony Diagnózy Lieky

Predikcia nákladov

Proces predikcie Výpočet počtu budúcich záznamov Predikcia budúcich záznamov Predikcia cenovej kategórie záznamu

Trénovanie, testovanie a validácia

Trénovanie a testovanie Validácia

Ďalšie kroky

Predikcia budúcich nákladov na pacienta

Cieľ práce

Dáta

Embedding

Diagnózy Lieky

Predikcia nákladov

Proces predikcie Výpočet počtu budúcich

záznamov Predikcia budúcich

Predikcia cenovej kategóri záznamu

testovanie a validácia

Trénovanie a testovanie

Výkony

Language-agnostic BERT sentence embedding model (LaBSE)



Lemmatizer + Word2vec model



Predikcia budúcich nákladov na pacienta

Cieľ práce

Dáta

Embeddi

Výkony

Liekv

بماد ماناه می شود المعالی

Proces prodikcie

Proces predikcie Výpočet počtu budúcici

Predikcia budúcich ráznamov

Predikcia cenovej kategório záznamu

Trénovanie, testovanie a validácia

Trénovanie a testova Validácia

Ďalšie kroky

Daisie Kroky



Diagnózy

▶ ICD-10-CM kódy

[znak][dvojčíslie].[dvojčíslie]

- J Choroby dýchacej sústavy
- J3 Iné choroby dýchacích ciest
- J38 Choroba hlasiviek a hrtana
- J38.0 Obrna hlasiviek a hrtana
- J38.01 Jednostranná čiastočná obrna hlasiviek a hrtana



Predikcia budúcich nákladov na pacienta

Sieľ práce

Jata

mbedding ^{(ýkony}

Diagnózy

Predikcia nákladov

Výpočet počtu

záznamov

Predikcia budúcich ráznamov

znamu Snovanie

testovanie, validácia

Trénovanie a testovan



Lieky

ATC kód

[znak][dvojčísle][dvojica znakov][dvojčíslie]

- N Centrálna nervová sústava
- N02 Analgetiká
- N02B Iné analgetiká a antipyretiká
- N02BA Kyselina salicylová a deriváty
- N02BA01 Acylpyrín

Predikcia budúcich nákladov na pacienta

Cieľ práce

Jata

Embedo

Výkony

Liekv

Prodikcia náklado

Proces predikcie

Výpočet počtu budú záznamov

Predikcia budúcich záznamov

Predikcia cenovej kategórie ráznamu

Trénovanie, testovanie a

validácia

Validácia



Cieľ práce

Dáta

Embedding

Výkony Diagnózy Lieky

Predikcia nákladov

Proces predikcie Výpočet počtu budúcich záznamov Predikcia budúcich záznamov Predikcia cenovej kategórie záznamu

Trénovanie, testovanie a validácia

Trénovanie a testovanie Validácia

Ďalšie kroky

Predikcia budúcich nákladov na pacienta

Cieľ práce

Dáta

Embedo

/ýkony

Diagnózy Lieky

Predikcia nákladov

Proces predikcie Výpočet počtu budúcich záznamov

záznamov Predikcia budúcich

Predikcia cenovej katego záznamu

Trénovanie, estovanie a

Trénovanie a testovanie



Proces predikcie

- 1. načítanie dát pacienta, embedding a normalizácia
- výpočet počtu záznamov na budúci rok
- predikcia budúcich záznamov
- 4. predikcia cenových kategórii budúcich záznamov
- výpočet celkovej cenovej kategórie

Predikcia budúcich nákladov na pacienta

Proces predikcie



Výpočet počtu budúcich záznamov

- Možnosť A) lineárna regresia počtu záznamov z predchádzajúcich rokov
- Možnosť B) generovanie nových záznamov kým generovaný timestamp nepresiahne rok od posledného skutočného záznamu

Predikcia budúcich nákladov na pacienta

Cieľ práce

Data

Embedd

Diagnózy

Predikcia nákladov

Proces predikcie

Výpočet počtu budúcich

Výpočet počtu budúcich záznamov

záznamov
Predikcia cenovej kategóri

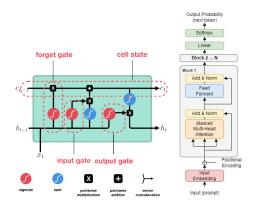
rénovanie, estovanie a

/alidácia Trénovanie a testova

Validácia

Predikcia budúcich záznamov

- využitie fixného maximálneho okna pre umožnenie paralelizácie trénovania
- testované modely
 - Long short-term memory
 - Decoder-only transformaer



Predikcia budúcich nákladov na pacienta

ieľ práce

Jata

Embedo

Výkony Diagnózy

ieky

Proces predikcie

Výpočet počtu budú záznamov

Predikcia budúcich záznamov

áznamu

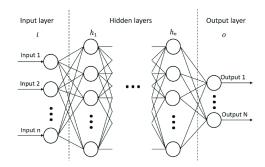
Trénovanie, testovanie a validácia

Trénovanie a testova Validácia



Predikcia cenovej kategórie záznamu

- namiesto konkrétnej ceny predikujeme cenovú kategóriu
- použitý model: Multilayer perceptron



Predikcia budúcich nákladov na pacienta

Predikcia cenovei kategórie záznamu

Cieľ práce

Dáta

Embedding

Výkony Diagnózy Lieky

Predikcia nákladov

Proces predikcie Výpočet počtu budúcich záznamov Predikcia budúcich záznamov Predikcia cenovej kategórie záznamu

Trénovanie, testovanie a validácia

Trénovanie a testovanie Validácia

Ďalšie kroky

Predikcia budúcich nákladov na pacienta

ieľ práce

Dáta

Embed

Výkony Diagnózy

Lieky

Predikcia nákladov

Proces predikcie Výpočet počtu budúci

záznamov

Predikcia budúcich

záznamov Predikcia cenovej kategó

Trénovanie, testovanie a validácia

Trénovanie a testovanie

Trénovanie a testovanie

- predikčné modely sa trénujú a testujú samostatne
- ako chyba predikcie záznamu sa využíva váhovaná euklidovská vzdialenosť od reálneho záznamu
- ako chyba predikcie cenovej kategórie sa využíva euklidovská vzdialenosť od one-hot vektora skutočnej skutočnej cenovej kategórie

Predikcia budúcich nákladov na pacienta

Trénovanie a testovanie

Validácia

- rozdelenie záznamov pacienta na posledný rok a staršie
- výpočet skutočnej celkovej cenovej kategórie pacienta v poslednom roku
- 3. predikcia počtu záznamov, jednotlivých záznamov, ich cenovej kategórie a celkovej cenovej kategórie
- výpočet chyby ako euklidovská vzdialenosť medzi predikovanou a skutočnou cenovou kategóriou v poslednom roku

Predikcia budúcich nákladov na pacienta

Cieľ práce

Jata

Embedo

Výkony Diagnózy

Dundikain méklandar

Proces predikcie

Výpočet počtu bu záznamov

> Predikcia budúcich ráznamov

redikcia cenovej kategó áznamu

Trénovanie, testovanie a validácia

rénovanie a testovanie

Validácia

Cieľ práce

Dáta

Embedding

Výkony Diagnózy Lieky

Predikcia nákladov

Proces predikcie Výpočet počtu budúcich záznamov Predikcia budúcich záznamov Predikcia cenovej kategórie záznamu

Trénovanie, testovanie a validácia

Trénovanie a testovanie Validácia

Ďalšie kroky

Predikcia budúcich nákladov na pacienta

ieľ práce

Julia

Embedding

Diagnózy

Prodikcia nákladov

Proces predikcie Výpočet počtu budúcich

záznamov

Predikcia budúcich

zaznamov Predikcia cenovej kategór

énovanie

testovanie a validácia

Trénovanie a testovanie Validácia

Ďalšie kroky

- trénovanie, testovanie a nastavovanie hyperparametrov modelov na väčšom datasete serveri
- výber najlepších hyperparametrov
- spísanie výsledkov

Predikcia budúcich nákladov na pacienta

