TECHNICKÁ UNIVERZITA V KOŠICIACH FAKULTA ELEKTROTECHNIKY A INFORMATIKY

ZÁKLADY CLOUDOVÝCH TECHNOLÓGIÍ

Zadanie č.2 – Prevod posunkovej reči na text

Skupina - Porsche 911 GT3 RS

TECHNICKÁ UNIVERZITA V KOŠICIACH FAKULTA ELEKTROTECHNIKY A INFORMATIKY

ZÁKLADY CLOUDOVÝCH TECHNOLÓGIÍ Zadanie č.2 – Prevod posunkovej reči na text Skupina - Porsche 911 GT3 RS

Študijný program: Inteligentné systémy

Študijný odbor: Informatika

Školiace pracovisko: Katedra kybernetiky a umelej inteligencie (KKUI)

Školiteľ: my

Konzultant: Ing. Ľubomír Urblík

ChatGPT, Gemini

Júlia Sičáková Pavol Slivoň Marián Hriň Filip Šajna

2024 Košice

Abstrakt v SJ

Táto práca sa zaoberá implementáciou a integráciou cloudových systémov na predikciu amerického znakového jazyka (ASL) a prekladu znakov do textu. Práca kombinuje techniky strojového učenia a cloudových technológií na vytvorenie efektívneho a dostupného riešenia. Model pre predikciu ASL bol natrénovaný lokálne a nasadený na platforme AzureML, čo umožňuje rýchle a presné rozpoznávanie gest. Webová aplikácia, ktorá využíva túto predikciu, je hostovaná na platforme Google Cloud Run, čo zabezpečuje škálovateľnosť a vysokú dostupnosť. Okrem toho, databáza pre užívateľské účty a registráciu je uložená na platforme AWS, čo zabezpečuje bezpečnosť a spoľahlivosť údajov. Toto riešenie ponúka užívateľom efektívny a spoľahlivý spôsob komunikácie prostredníctvom znakového jazyka, s využitím moderných cloudových technológií.

Kľúčové slová v SJ

Cloudové systémy, predikcia ASL, preklad znakového jazyka, AzureML, Google Cloud Run, AWS, strojové učenie, webová aplikácia, databáza

Abstract

This assignment focuses on the implementation and integration of cloud systems for predicting American Sign Language (ASL) and translating sign language into text. The work combines machine learning techniques and cloud technologies to create an efficient and accessible solution. The ASL prediction model was trained locally and deployed on the AzureML platform, enabling fast and accurate gesture recognition. The web application utilizing this prediction is hosted on the Google Cloud Run platform, ensuring scalability and high availability. Additionally, the database for user accounts and registration is stored on the AWS platform, ensuring data security and reliability. This solution offers users an efficient and reliable means of communication through sign language, utilizing modern cloud technologies.

Keywords

Cloud systems, ASL prediction, sign language translation, AzureML, Google Cloud Run, AWS, machine learning, web application, database

Zadanie práce

Zadanie 2 - Aplikácia s podporou cloud-u (IS + HI)

Zadanie pre študentov HI a IS bude mať nasledujúce znenie:

Vytvorte webovú alebo mobilnú aplikáciu, ktorá bude riešiť zadanú problematiku. Táto aplikácia by mala obsahovať klasicky backend a frontend, nasadená by mala byť na cloud-e jedného z veľkých providerov (prednáška 3). Zároveň by pre zabezpečenie svojej funkcionality mala využívať cloudové služby, ktoré bude poskytovať aspoň 1 provider, ktorý je iný ako provider poskytujúci hosting. Na ukladanie dát v aplikácii je potrebné použiť databázu, čo budete ukladať necháme na Vás, no niečo áno, a na komunikáciu so službami API poskytnuté providermi. Aplikáciu je postačujúce nasadzovať na úrovni PaaS, nie je potrebné fungovať na úrovni laaS. Zadanie bude určené pre skupinu 4 študentov. Na dosiahnutie plného počtu bodov členom skupiny, bude potrebná jeho aktívna účasť na zadaní, ktorá bude kontrolovaná pomocou systému GitLab alebo GitHub. Každý tím bude disponovať privátnym Git repozitárom, na ktorý bude mať prístup ich cvičiaci. Jednotlivé úlohy pre jednotlivých členov team-u by mali byť rozbité do To Do listu a tieto by mali byť adresované jednotlivými commitmi. Voľba technológie, platformy na hosting, rozloženia UI a použitých služieb je plne vo Vašej réžií. Na tvorbu týchto zadaní by Vám mali stačiť študentské/trial verzie od cloudových providerov.



Poďakovanie Radi by sme sa poďakovali všetkým, ktorí prispeli k úspešnému dokončeniu tejto práce. Vďaka ich podpore a úsiliu sme boli schopní implementovať a integrovať cloudové systémy na predikciu amerického znakového jazyka a preklad znakov do textu. Bez ich podpory by tento projekt nebol možný. Osobitné poďakovanie patrí nášmu vedúcemu práce za jeho cenné rady a usmernenie počas celej práce. Tiež by sme chceli poďakovať sami sebe za úspešné dokončenie zadania aj keď sa to nezaobišlo bez psychickej bolesti a beznádeje.

Obsah

5.1. Marián Hriň	15
5.2. Filip Šajna	16
5.2.1. Program	
5.2.2. Back-end:	
5.3. Pavol Slivoň	
5.4. Júlia Sičáková	
5.4.1. Farebná paleta	19
5.4.2. Úvodná strana	
5.4.3. Login page	
5.4.4. Stránka pre registráciu	

Úvod

V súčasnej digitálnej ére, kde technologický pokrok formuje spôsob, ako komunikujeme a interagujeme s okolitým svetom, je dôležité, aby sme zabezpečili, že všetci jednotlivci majú rovnaké možnosti pre vyjadrovanie a porozumenie. Americký znakový jazyk (ASL) je dôležitým nástrojom komunikácie pre komunitu osôb so sluchovým postihnutím, a preto je nevyhnutné poskytnúť efektívne technologické riešenia na podporu jeho používania. V tejto práci sa zaoberáme implementáciou a integráciou cloudových systémov na predikciu ASL gest a preklad znakov do textu. Tento projekt spočíva v kombinácii pokročilých techník strojového učenia s využitím vysoko dostupných cloudových platform, ako je AzureML, Google Cloud Run a AWS, aby sme vytvorili robustné a efektívne riešenie. Cieľom tejto práce je vytvoriť nástroj, ktorý umožní ľuďom so sluchovým postihnutím jednoduchšie a efektívnejšie komunikovať prostredníctvom znakového jazyka, pričom využíva výhody moderných technologických riešení dostupných v cloude. Tento úvod poskytuje prehľad o cieľoch a metódach tejto práce, ako aj o dôležitosti poskytovania inkluzívnych technologických nástrojov pre všetkých používateľov.

1. Formulácia úlohy a cieľ práce

Cieľom tejto práce je implementovať a integrovať cloudové systémy na predikciu amerického znakového jazyka (ASL) a preklad znakov do textu. Konkrétne sa zameriame na natrénovanie modelu strojového učenia na rozpoznávanie ASL gest, jeho nasadenie na platforme AzureML a vytvorenie webovej aplikácie, ktorá využije tento model na poskytovanie prekladu znakov do textu. Webová aplikácia bude nasadená na platforme Google Cloud Run a bude integrovaná s databázou na platforme AWS na správu užívateľských účtov a registráciu. Cieľom je poskytnúť užívateľom efektívne a spoľahlivé prostriedky na komunikáciu prostredníctvom znakového jazyka pomocou moderných cloudových technológií.

2. Teoretický rozbor zvolenej témy

Americký znakový jazyk (ASL) je kompletný jazykový systém, ktorý sa používa na komunikáciu medzi osobami so sluchovým postihnutím. ASL sa skladá z vizuálnych gest a pohybov rúk a tela, pričom každé gest reprezentuje konkrétny význam alebo koncept. Rozpoznávanie ASL gest je náročný problém v oblasti strojového videnia, pretože gestá môžu mať rôzne tvary, polohy a rýchlosti. Na riešenie tohto problému sa využíva strojové učenie, konkrétne metódy hlbokého učenia (deep learning), ktoré sú schopné naučiť sa zložité vzory a štruktúry v dátach. Modely hlbokého učenia, ako sú konvolučné neurónové siete (CNN), sú efektívne v rozpoznávaní obrazových vzorov a môžu byť úspešne použité na rozpoznávanie ASL gest.

Cloudové systémy poskytujú ideálne prostredie pre nasadenie a integráciu modelov strojového učenia. Služby ako AzureML, Google Cloud AI a AWS umožňujú jednoduché natrénovanie a nasadenie modelov strojového učenia a poskytujú škálovateľné a vysoko dostupné prostredie pre ich prevádzku. Google Cloud Run a podobné služby poskytujú platformy pre nasadenie a prevádzku webových aplikácií, ktoré využívajú modely strojového učenia. Tieto platformy zabezpečujú flexibilitu, škálovateľnosť a vysokú dostupnosť aplikácií.

Integrácia s databázovými systémami, ako je AWS DynamoDB alebo Google Cloud Firestore, umožňuje ukladanie a správu užívateľských účtov a dát používaných webovou aplikáciou na spracovanie ASL gest a preklad do textu. Spoločná práca týchto technológií umožňuje vytváranie efektívnych a spoľahlivých systémov pre rozpoznávanie ASL gest a preklad do textu, čím poskytuje osobám so sluchovým postihnutím moderné a inkluzívne nástroje komunikácie.

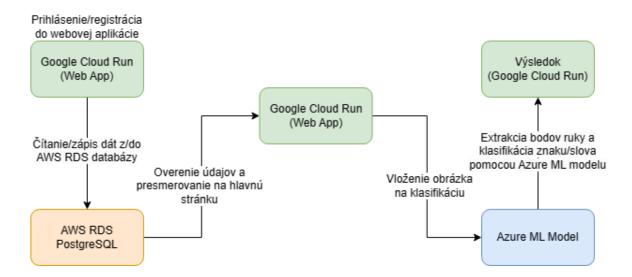
3. Analýza stavu problematiky

Vstupné dáta modelu pre trénovanie sú fotky dlane, ktoré ukazujú jednotlivé znaky a k ním príslušné labels. Na trénovanie modelu bola použitá neurónová sieť. Hlavným problémom bolo vybrať vhodného poskytovateľa cloudových služieb.

4. Návrh a implementácia riešenia zvolenej problematiky

- ☑ Vytvorenie a natrénovanie modelu pre klasifikáciu
- ☑ Vytvorenie AWS RDS PostgreSQL databázy
- ☑ Nahranie modelu do Azure ML a vytvorenie deployment-u pre model
- ✓ UI|UX webovej aplikácie
- ☑ Back-end webovej aplikácie
- ☑ Dockerizácia/kontajnerizácia webovej aplikácie
- ☑ Vytvorenie spustiteľnej inŠtancie webovej aplikácie na Google Cloud Run

4.1. Diagram použitých služieb a ich prepojenia



5. Príspevok jednotlivých členov tímu k funkčnosti riešenia

5.1.Marián Hriň

5.1.1. Vytvorenie dátovej množiny

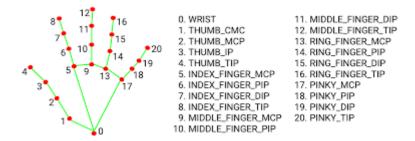
Prvým krokom bolo vytvorenie dátovej množiny. Anglická abeceda má 26 znakov. Ďalších 5 znakov, poprípade fráz bolo:

- Space: na medzery v texte,
- Delete: ak sa daný používateľ pomýlil v skladbe vety, môže posledný znak vymazať,
- ILY: fráza I love you,
- Hello: pre rýchle pozdravenie,
- PEWPEW: pri tvorení dátovej množiny som sa nudil.

Dátová množina dokopy obsahuje 7750 obrázkov, Čiže jeden znak má 250 obrázkov s rôznou kompozíciou pre zvýšenie presnosti modelu.

5.1.2 Kľúčové body na dlani

Po úspešnom nazbieraní fotiek, nasledovalo uloženie takzvaných landmarks do pickle súboru. Landmarky je možné vidieť na obrázku nižšie (červené body).



5.1.3 Trénovanie modelu

Bolo netrénovaných viacero modelov, napríklad: Decision Tree, neurónová sieť pomocou TensorFlow-u, taktiež pomocou PyTorch-u, XGBoost, AdaBoost a v neposlednom rade Random Forest. Podľa metriky accuracy väčšina týchto modelov vykazovalo presnosť 95-100%, ale pri realtime testovaní tieto modely zlyhávali (predikovali zlé písmená). Jediný dobrý model, ktorý fungoval aj realtime bol Random Forest. Ten sme nakoniec aj použili. Model bol uložený ako pickle súbor a nahraný do MS AzureML.

5.2.Filip Šajna

5.2.1. Program

Dockerizácia front-endu a back-endu do Docker Image

Docker je nástroj na kontajnerizáciu, ktorý umožňuje spúšťať aplikácie v izolovaných prostrediach nazývaných kontajnery. Tento proces zjednodušuje vývoj, nasadenie a prevádzku softvérových aplikácií tým, že ich balíčkuje so všetkými potrebnými závislosťami.

Vytvorenie Dockerfile:

Dockerfile je textový súbor obsahujúci príkazy na zostavenie Docker image. Tento súbor popisuje konfiguráciu a prostredie, v ktorom bude aplikácia spúšťaná.

5.2.2. Back-end:

Back-end časť aplikácie sa zameriava na spracovanie dát a logiku, ktorá nie je priamo prístupná používateľom.

• Vytvorenie funkcionality pre registráciu a prihlasovanie:

Implementácia procesu registrácie a prihlásenia používateľov je kľúčová pre správne fungovanie webovej aplikácie. To zahŕňa validáciu, ukladanie a overovanie používateľských údajov.

Zápis údajov do databázy na AWS:

Údaje, ktoré používateľ zadá pri registrácii, sú ukladané do databázy na cloudovej službe AWS RDS, konkrétne do PostgreSQL databázy.

Routing a správa prístupov:

Definovanie a riadenie navigácie v rámci webovej aplikácie, vrátane zabezpečenia prístupu používateľov k jednotlivým stránkam alebo podstránkam.

Azure:

Azure je cloudová platforma od spoločnosti Microsoft, ktorá poskytuje škálovateľné riešenia pre ukladanie, spracovanie a analyzovanie dát.

Vytvorenie služby v Azure Machine Learning workspace:

Vytvorenie pracovnej plochy v Azure Machine Learning, ktorá umožňuje vývoj, správu a nasadzovanie modelov strojového učenia.

• Nahrávanie modelov do Azure ML Studio:

Modely vyvinuté v rôznych formátoch (.pkt) sú nahrávané do Azure ML Studio, kde môžu byť ďalej spracované a nasadené.

Vytvorenie skriptu pre Azure ML Endpoint:

Vytvorenie skriptu v jazyku Python pre definovanie správania sa endpointu, ktorý poskytuje prístup k modelom strojového učenia.

• Vytvorenie custom environmentu pre deployment modelu:

Vytvorenie vlastného prostredia (environment) pre nasadenie modelov na základe špecifických požiadaviek knižnice scikit-learn verzie 1.1.

• Vytvorenie deploymentu a endpointu pre model:

Nasadenie modelu a definovanie endpointu, ktorý umožňuje prístup k jeho funkciám. Nastavenie výpočtových zdrojov (PC) pre optimálne spracovanie požiadaviek.

AWS:

Amazon Web Services (AWS) je poskytovateľ cloudových služieb, ktorý ponúka širokú škálu služieb pre ukladanie dát, vývoj aplikácií a správu infraštruktúry.

• Vytvorenie a sprístupnenie PostgreSQL databázy:

Vytvorenie a konfigurácia relačnej databázy PostgreSQL na AWS RDS (Relational Database Service) a definovanie prístupových práv pre vonkajšie pripojenia.

Vytvorenie bezpečnostných pravidiel pre databázu:

Nastavenie bezpečnostných pravidiel pre obmedzenie prístupu k databáze a zabezpečenie dát.

Vytvorenie tabuľky pre ukladanie údajov používateľov:

Definovanie štruktúry tabuľky v databáze, do ktorej budú ukladané registračné a prihlasovacie údaje používateľov.

Google Cloud:

Google Cloud je cloudová platforma od spoločnosti Google, ktorá poskytuje širokú škálu služieb pre ukladanie, spracovanie a nasadzovanie aplikácií.

Nahratie Docker Image do Google Container Registry:

Nahratie Docker image obsahujúceho finálnu aplikáciu do Google Container Registry, ktorý slúži na ukladanie a správu kontajnerov.

• Deploy aplikácie cez Google Cloud Run:

Nasadenie finálnej aplikácie do služby Google Cloud Run, ktorá umožňuje spúšťať kontajnery v cloudovom prostredí. Nastavenie smerovania a prístupových práv k aplikácii.

5.3. Pavol Slivoň

Vytvorenie a uvedenie do funkčnosti proces videa - to, ako sa posiela obrázok z cloudu do Pythonu a následne z Pythonu do webu.

Úprava HTML, CSS a pridanie ciest na presmerovanie používateľa na nasledujúce strany.

Priebežné riešenie a upravovanie back-end časti zadania, debugging.

5.4. Júlia Sičáková

Vytvorenie front-end časti - dizajnu webovej stránky pomocou HTML, CSS, Javascriptu a Figmy. Cieľom vytvárania stránky bolo použíť také farby aby boli používateľovi príjemné a vytvoriť také prostredie, ktoré bude pre používateľa intuitívne. Dizajn na všetkých stránkach je konzistentný, zaoblenie rohov boxov, použitie tých istých farieb elementov, tlačidlá majú tie isté farby aj veľkosť. Použitie zaoblených elementov malo za cieľ spríjemniť používateľovi prechod aplikáciou.

5.4.1. Farebná paleta

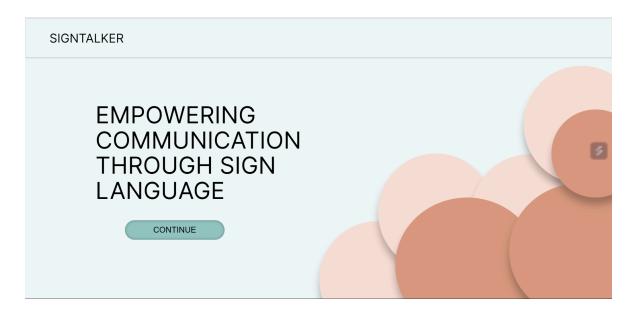


Farby pre dizajn webovej stránky bol vybraný s ohľadom na používateľa a preto sú farby tlmenejšie a navodzujú v používateľovi pocit pokoja.

5.4.2. Úvodná strana

Prvá a zároveň úvodná strana landing_page.html obsahuje:

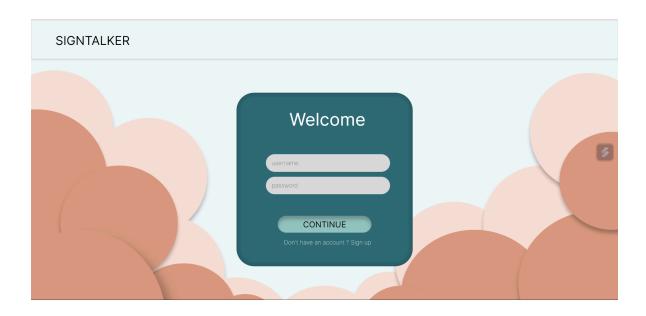
- úvodný text na privítanie používateľa
- logo SIGNTALKER, ktoré je interaktívne a po kliknutí používateľa presmeruje na landing page
- CONTINUE button ktorý presmeruje používateľa na nasledujúcu stranu
- elementy s animáciami, ktoré dotvárajú dizajn stránky



5.4.3. Login page

Login page obsahuje:

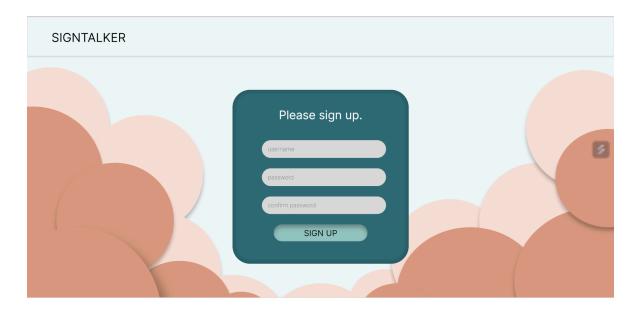
- prihlasovací formulár kde používateľ má možnosť sa prihlásiť pre prístup k ďalšej časti webovej aplikácie
- možnosť zaregistrovať sa linkom "Don´t have an account? Sign up", ktorý po kliknutí naň
 presmeruje používateľa na stránku pre registráciu
- v ľavom hornom rohu sa nachádza logo "SIGNTALKER", ktoré po kliknutí presmeruje používateľa na úvodnú stránku



5.4.4. Stránka pre registráciu

Na tejto stránke sa nachádza:

- registračný formulár, pričom heslo sa používateľovi zobrazuje len ako bodky pre vyššiu bezpečnosť
- "SIGNUP" button, ktorý po úspešnej registrácii používateľa presmeruje na hlavnú stránku
- v ľavom hornom rohu sa nachádza "SIGNTALKER" logo, ktoré po kliknutí naň presmeruje používateľa na úvodnú obrazovku

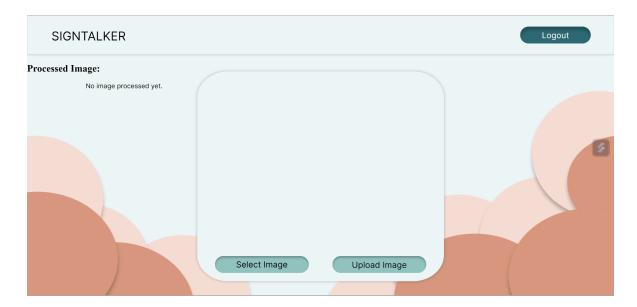


5.4.5. Hlavná stránka

Na hlavnej stránke sa nachádza:

box, kde sa zobrazuje nahraný obrázok

- v boxe sú dve tlačidlá, "Select Image", ktorým používateľ môže zo zariadenia vybrať požadovaný obrázok s dlaňou znázorňujúci nejaký ASL znak a "Upload Image" button, ktorým používateľ nahrá daný obrázok do modelu
- výstup, teda obrázok dlane už so zisteným znakom sa zobrazí v tom istom boxe
- v ľavom hornom rohu sa tak ako na ostatných stránkach nachádza "SIGNTALKER" logo,
 ktoré po kliknutí presmeruje používateľa na úvodnú stranku
- v pravom hornom rohu sa nachádza "Logout" button, ktorý po kliknutí odhlási používateľa
 a presmeruje ho na prihlasovaciu stránku



Záver

Stanovené body zadania sa nám podarilo úspešne splniť. V zadaní sa nachádzajú všetky potrebné komponenty, ktoré boli vo formulácií zadania teda obsahuje front-end, back-end, aplikácia je nasadená na cloud-e, využíva cloudové služby. Prebehol aj testovanie presnosti modelu na predikciu jednotlivých znakov, pričom sme dosiahli takmer 100% úspešnosť.

Používateľská príručka

- Aplikácia sa spúšťa pomocou tlačidla Start v Google Cloud Run alebo pomocou app.py v priečinku zadania
- 2. Následne sa používateľovi zobrazí úvodná obrazovka webovej aplikácie
- 3. Po stlačení tlačidla "CONTINUE" na úvodnej obrazovke je používateľ presmerovaný na stránku pre prihlásenie Login page kde je požiadaný aby zadal prihlasovacie meno "username" a heslo "password".
- 4. Po úspešnom prihlásení sa používateľovi zobrazí hlavná stránka aj s prázdnym oknom kde bude nahratý obrázok a s tlačidlami pre výber obrázka - "Select Image" a nahratie obrázku - "Upload Image"
- 5. Po stlačení "Upload Image" tlačidla sa obrázok spracuje a následne sa zobrazí obrázok aj s predikovaným znakom.
- 6. Po ukončení používania webovej aplikácie môže používateľ stlačiť "Log out" button, ktorý ho odhlási a presmeruje znova na Log in stránku.