

Učení neuronové sítě pro klasifikaci záměru hlasovým dialogem

Vladimíra Kimlová¹

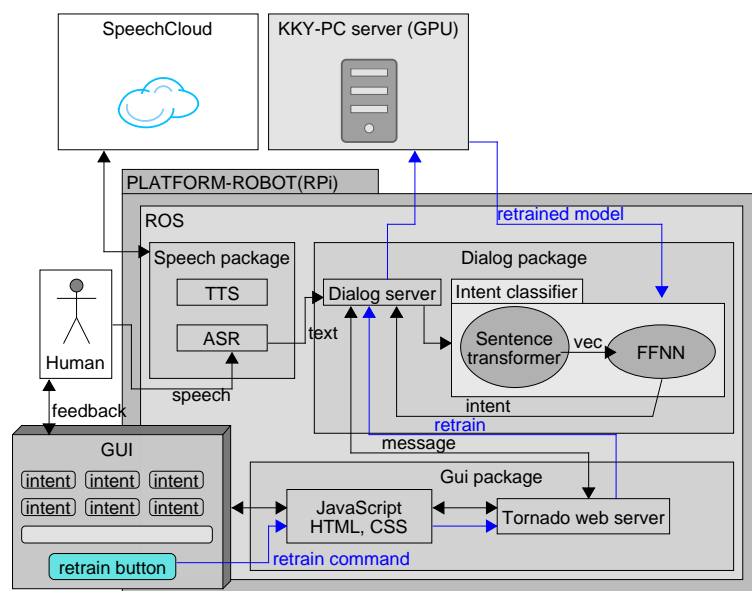
1 Motivace

Neuronové sítě jsou dnes dominantním nástrojem pro mnoho úloh v oblasti strojového učení a umělé inteligence. Avšak pro dosažení vysoké přesnosti řešení dané úlohy je potřeba velké množství anotovaných dat pro trénování sítě, což může být časově náročné a drahé.

2 Zpětnovazební hlasové učení

V rámci této práce byla vyvinuta metoda trénování neuronové sítě, která umožňuje lidskou interakci s trénovacím algoritmem v reálném čase. Tato metoda byla demonstrována na jednoduché dopředné síti pro klasifikaci záměru, která byla optimalizována pro použití na Raspberry Pi. Pro zprostředkování komunikace s uživatelem byla využita katedrální platforma *Speechcloud*. Kromě toho byla využita moderní technologie *Sentence Transformer* sloužící k převodu rozpoznávaného vstupního textu na vektory reprezentující sémantickou informaci. Proces postupného učení umožňuje síti zlepšovat své klasifikační schopnosti a přizpůsobit se potřebám uživatele. Přeučení sítě se uskutečňuje velice rychle díky transferu parametrů z již předtrénované neuronové sítě na robustním datasetu a *fine-tuningem* pouze poslední plně propojené vrstvy.

Diagram na Obrázku 1 naznačuje propojení jednotlivých stavebních bloků celé aplikace a tok informací mezi uživatelem a jednotlivými bloky. Běh všech skriptů v rámci Raspberry Pi 4B, je zajištěn pomocí softwarového frameworku ROS. Pro každou funkční část aplikace byl vytvořen ROS balík.



Obrázek 1: Propojení jednotlivých částí v rámci celé aplikace

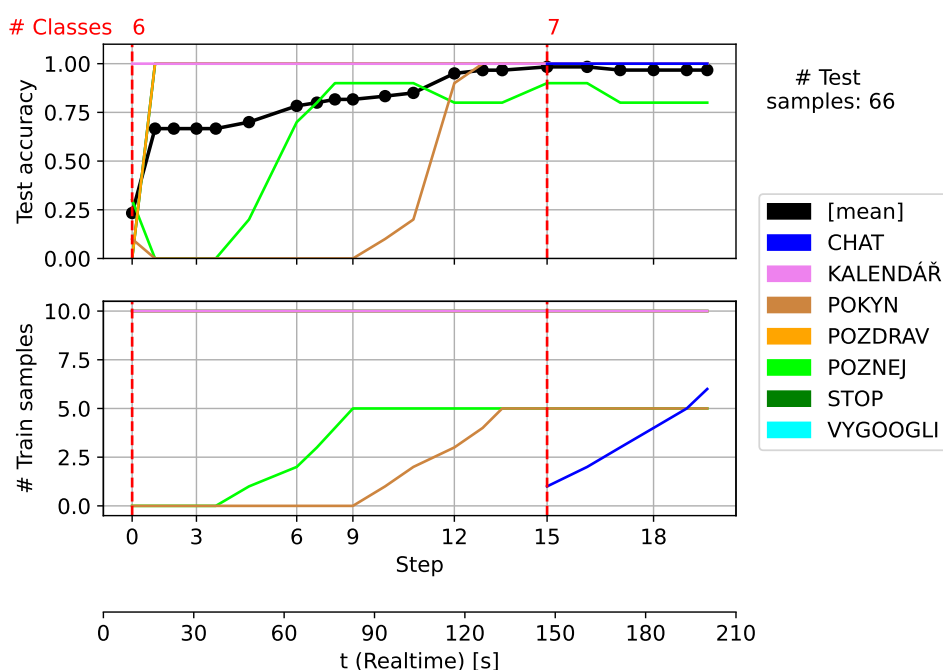
¹ studentka bakalářského studijního programu Kybernetika a řídicí technika, specializace Umělá inteligence a automatizace, e-mail: jivl@students.zcu.cz

3 Experimentální testování

Následující experiment simuluje situaci opakovaného přeučování modelu na základě zpětné vazby od uživatele. Testovací a validační množina obsahovala 6 tříd po 10 vzorcích, v rámci experimentu byly do trénovací množiny iterativně přidávány nové vzorky. Na konci každého kroku byl model přeučen.

Simulace proběhla pro scénář, kdy počáteční trénovací množina obsahovala 10 vzorků pro 4 vybrané třídy z testovací množiny. V prvních čtyřech krocích bylo přidáno po jednom vzorku do již naučených tříd (trénovací, validační ani testovací množina se tedy neměnila). V dalších krocích bylo přidáno po 5 vzorcích do tříd POZNEJ a POKYN. Nakonec byla zavedena ještě úplně nová třída CHAT a bylo do ní přidáno 6 vzorků.

Obrázek 2 znázorňuje iterativní přidávání vzorků v rámci výše zmíněného scénáře. Červenou barvou je znázorněn počet tříd v testovacích datech. První podgraf znázorňuje vývoj přesnosti na testovacích datech pro jednotlivé třídy, druhý podgraf zobrazuje počty trénovacích vzorků pro jednotlivé třídy. Vodorovná osa obou podgrafů zobrazuje počet aktuálních kroků – jejich rozložení je závislé na reálném čase na vodorovné ose v dolní části. Na základě obrázku lze identifikovat třídy, které nejvíce ovlivňují výslednou přesnost a zaměřit se na ně při dalším tréninku modelu. V tomto případě by to byla třída POZNEJ.



Obrázek 2: Vývoj přesnosti pro jednotlivé třídy při iterativním přidávání vzorků

Tento experiment simuloval reálnou situaci, kdy je zapotřebí doplnit model o nové třídy. S využitím předtrénované sítě lze takového rozšíření dosáhnout velmi rychle. Cílem je však zaručit, že se síť nesoustředí pouze na nové třídy a nezapomíná na již naučené informace (staré třídy). Důležitým přínosem přístupu *human-in-the-loop* je možnost sledovat konkrétní třídy, které ovlivňují úspěšnost klasifikace, a zaměřit se na ně v průběhu tréninku. Tento přístup poskytuje výhodu oproti opakovanému trénování na celém datasetu, protože umožňuje efektivnější vylepšování výkonu modelu.