

17

Odhad emocí z textu a analýza tweetů

Autor: Aneta Dufková

Vedoucí: Ing. Igor Szóke, Ph.D.

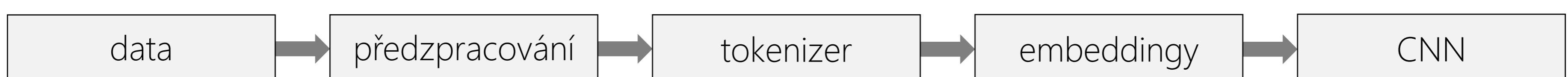


Motivace

Na Twitteru přibývají tisíce tweetů za minutu. Hromadné zpracování a vyhodnocení emocí skrývajících se v tweetech pomůže například v situacích, kdy:

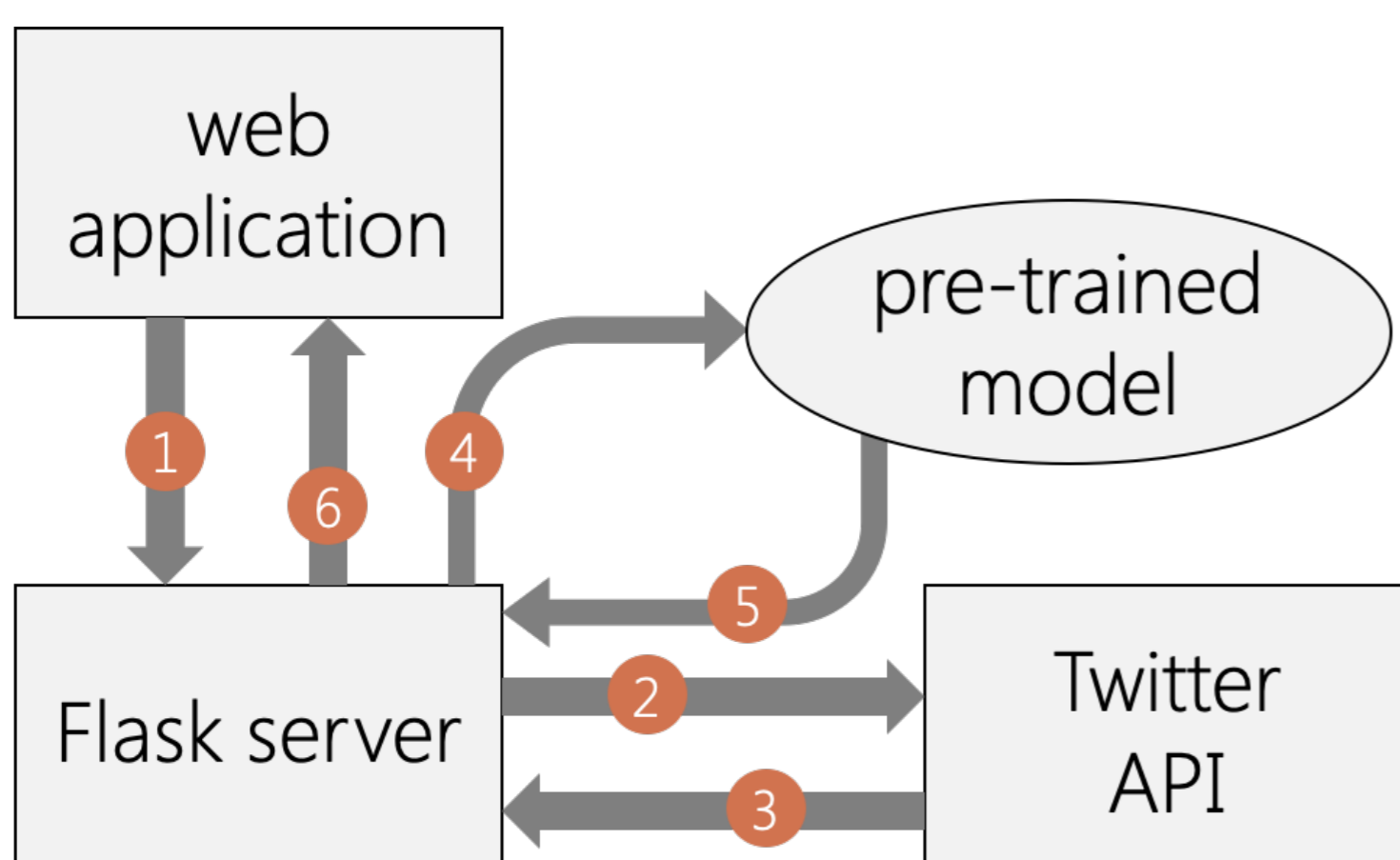
- firmy chtějí vědět, co si o nich myslí zákazníci,
- vládu zajímá, zda se obyvatelé bouří proti novému nařízení,
- zákazník se díky názorům ostatních rozhodne, jaký výrobek si koupí.

Řešení



Základem systému na rozpoznávání emocí z textu jsou vhodná data pro trénování. Tato data jsou předzpracována, pomocí tokenizeru se z nich vytvoří slovník a následně, s předtrénovaným modelem Word2Vec, embeddingy. Využita byla konvoluční síť s pěti konvolučními vrstvami. Výsledný model pro zadaný vstup určí pravděpodobnost každé ze tří tříd. Počáteční experimenty počítaly pouze s binární klasifikací, nicméně ukázalo se, že neutrální třída je důležitá už ve fázi trénování.

Využití



Výsledek ukazuje webová aplikace WhatDoesTwitterThink.com, která umožňuje zjistit, co si myslí uživatelé Twitteru o zadaném tématu. Aplikace je postavena na modelu klient-server, klienta představuje webové rozhraní, server byl vytvořen pomocí frameworku Flask.

Webová aplikace také sbírá zpětnou vazbu od uživatelů a ukládá ji do databáze. Uživatelé mohou nesprávně klasifikované tweety opravit a pomoci tak vytvořit nový dataset, který povede k přesnějším výsledkům.

WHAT
DOES
TWITTER
THINK
.com

