

Розробка автоматизованої підсистеми для оцінки правдоподібності інформації в мережі Інтернет із використанням нейронних мереж

ВИКОНАВ
СТУДЕНТ 4-ГО КУРСУ
ГРУПИ КА-33 ІПСА «КПІ»
МИРОНЕНКО ОЛЕКСІЙ
КЕРІВНИК
АСИСТЕНТ КУХАРЄВ С. А.

Вступ

Об'єкт дослідження: інформаційні ресурси мережі інтернет.

Предмет дослідження: застосування застосування методів машинного навчання, зокрема штучних нейронних мереж для класифікації текстової інформації.

Мета дослідження: розробка і практична реалізація алгоритму автоматичного визначення рівня правдоподібності текстової інформації, розміщеної в мережі інтернет.

Методи дослідження: методи обробки природної мови, класифікація за допомогою багат шарової моделі штучної мережі перцептрон.

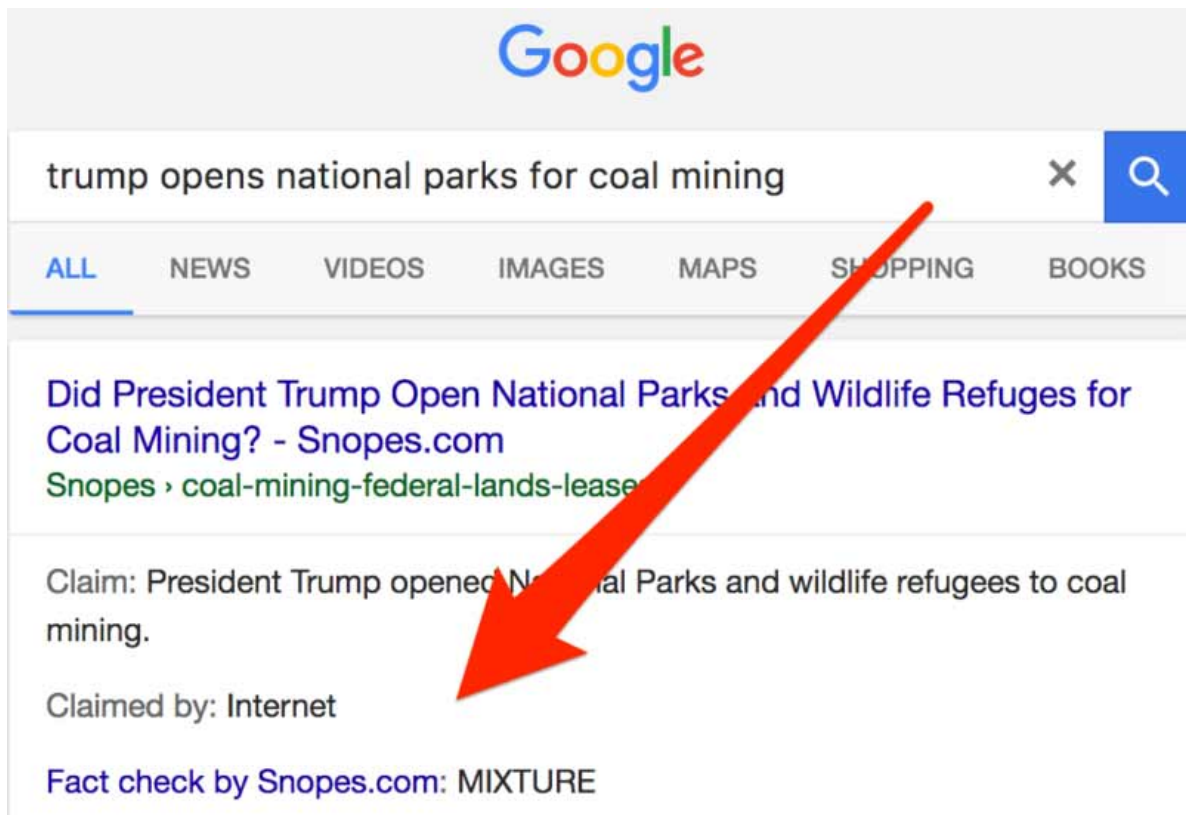
Актуальність

Сучасний інформаційний простір, зокрема Інтернет, крім виконання функцій обміну думками та отримання інформації його користувачами в період інформаційного протистояння, стає об'єктом і засобом інформаційного керування. Серед користувачів мережі з'являються групи людей або окремі особи, які навмисно поширюють помилкову або спотворену інформацію.

За таких умов користувачі мережі Інтернет мають критично підходити до переглянутої інформації, звертати увагу на велику кількість факторів для визначення її правдоподібності.

Огляд існуючих рішень

Google Fact Check



Google

trump opens national parks for coal mining

ALL NEWS VIDEOS IMAGES MAPS SHOPPING BOOKS

Did President Trump Open National Parks and Wildlife Refuges for Coal Mining? - Snopes.com
Snopes › coal-mining-federal-lands-lease

Claim: President Trump opened National Parks and wildlife refuges to coal mining.

Claimed by: Internet

Fact check by Snopes.com: MIXTURE

A large red arrow points from the search bar area down to the fact check result.

Fake News Detector AI



FAKE NEWS DETECTOR AI

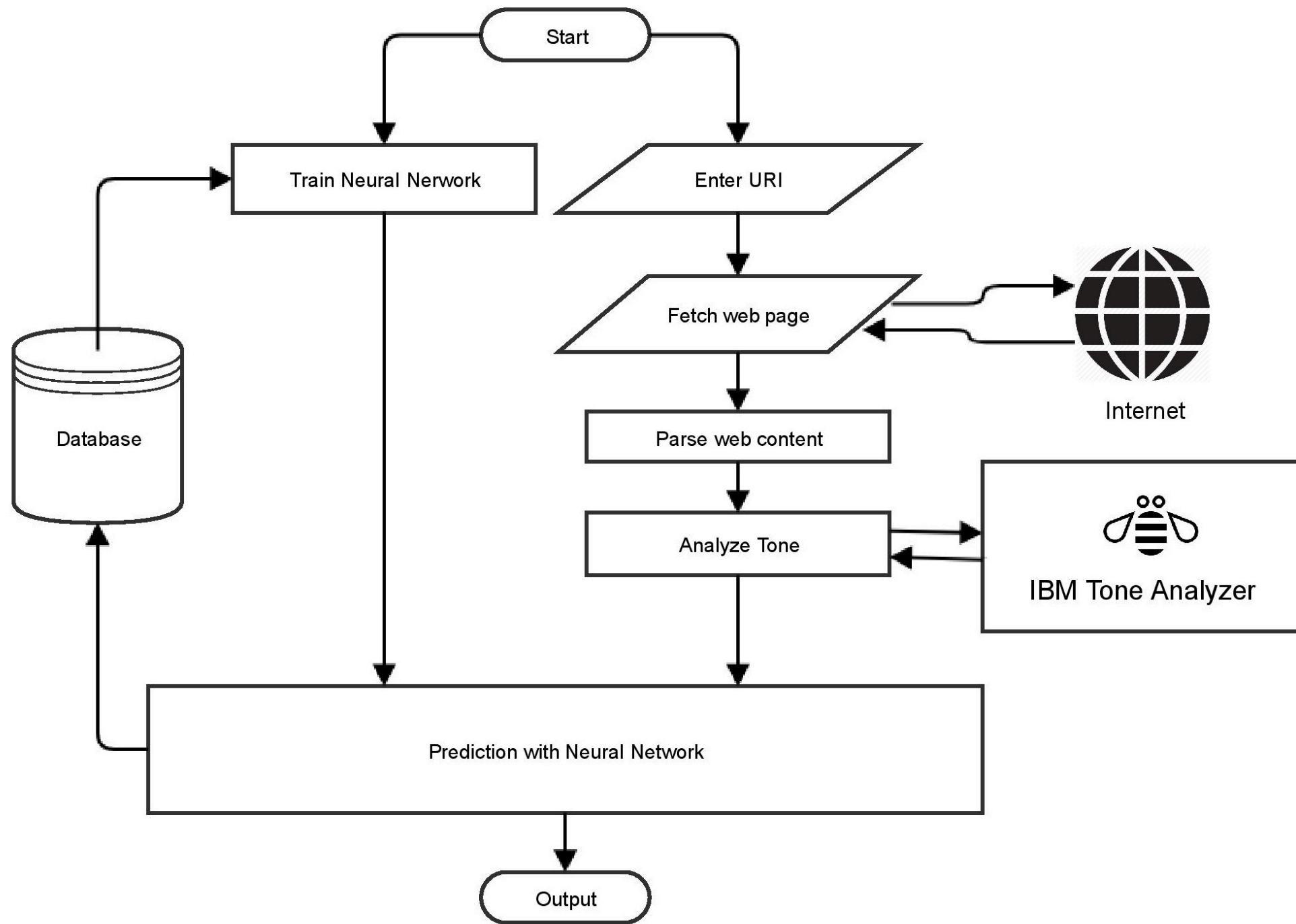
Ex: nytimes.com

Detect fake news sites using the power of artificial intelligence!

We analyze websites to see if they are similar to known fake news sites using a neural network. The same technology is used to power other artificial intelligence applications, like Siri and self-driving cars!

Created by [Karan Singhal](#). Download the [Chrome extension!](#)

Запропонований алгоритм системи



Опис вхідних даних

В прикладі розглядається стаття

<http://www.apfn.org/apfn/Skull & Bones.htm> про орден

Череп та кісток Єльського університету, розміщена на ресурсі www.apfn.org, який містить багато статей про теорії світової змови тощо.

В якості тренувальної вибірки беремо масив даних розмірності 100, у форматі ключ-значення, де ключем є URI (адреса) веб-сторінки, а значення – логічне 1 (правдивий ресурс) або 0 (ресурс не надійний)

The Order of the Skull and Bones

Everything you always wanted to know, but were afraid to ask.



BONESMEN



Практична реалізація

FakeCheck

Розпізнай підроблені новини разом з FakeCheck

Введіть адресу сайту нижче щоб перевірити його на правдоподібність

http://www.thisislocallondon.co.uk/news/15302696.19_people_die_in_suspected_terr

🔍 Проаналізувати

Про FakeCheck:

Що таке підроблені новини?

Підроблені новини це хибна інформація, яка подається читачеві як правдивий факт. Підроблені новини створюють під виглядом справжніх журналістських новин, але зі шкідливою метою. Їх часто використовують зі

Про FakeCheck

FakeCheck - система, що автоматично перевіряє, чи є дане джерело інформації підробкою. Вона використовує методи штучного інтелекту, зокрема аналіз тону повідомлення Tone Analyzer, і пропускає дані через треновану штучну нейронну систему для

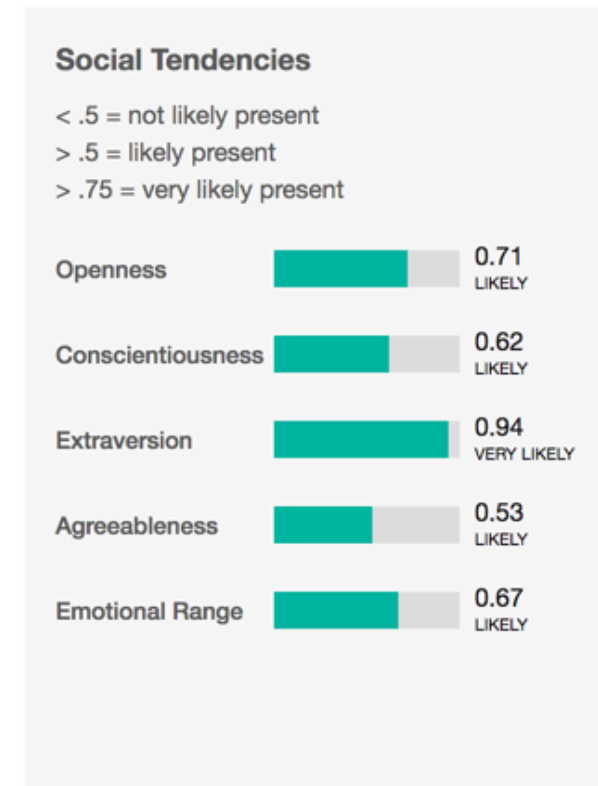
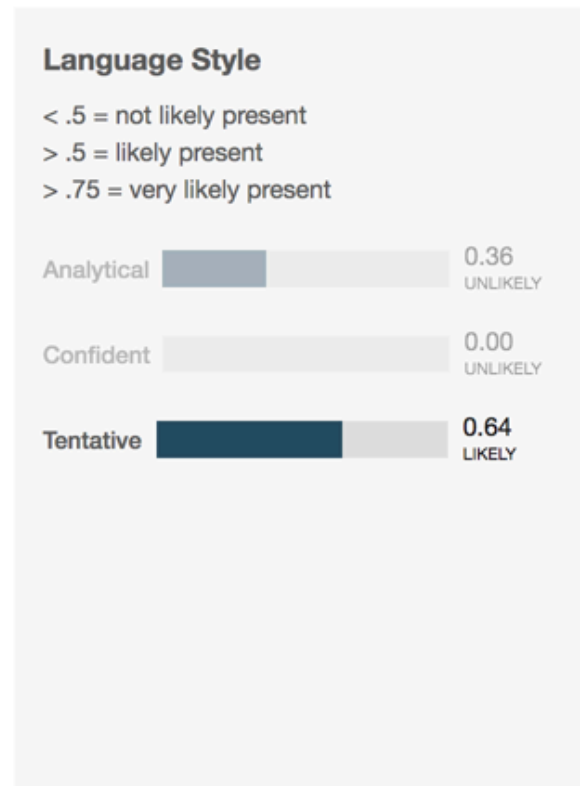
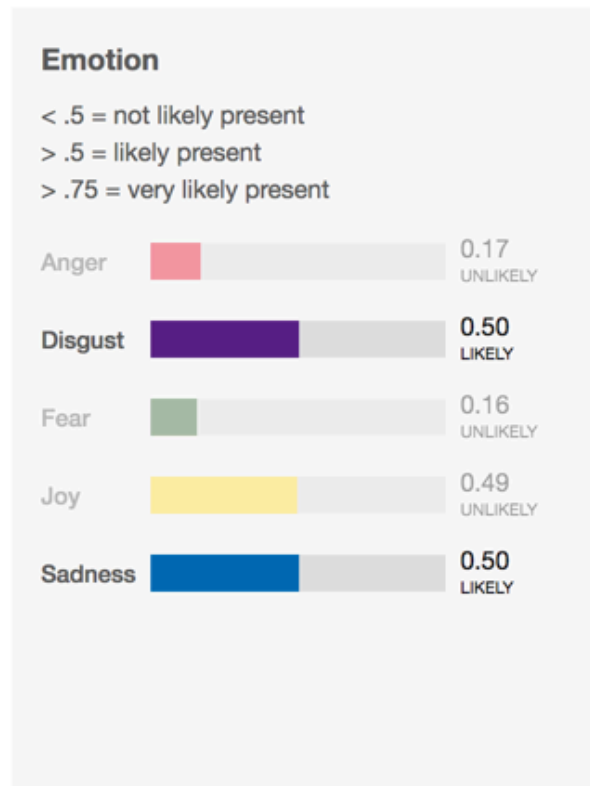
Як FakeCheck може Вам допомогти?

Будь-хто може протидіяти поширенню фіктивних новин притримуючись деяких простих правил.

Слідкуйте за тим, чим ви ділитесь! Не

Практична реалізація

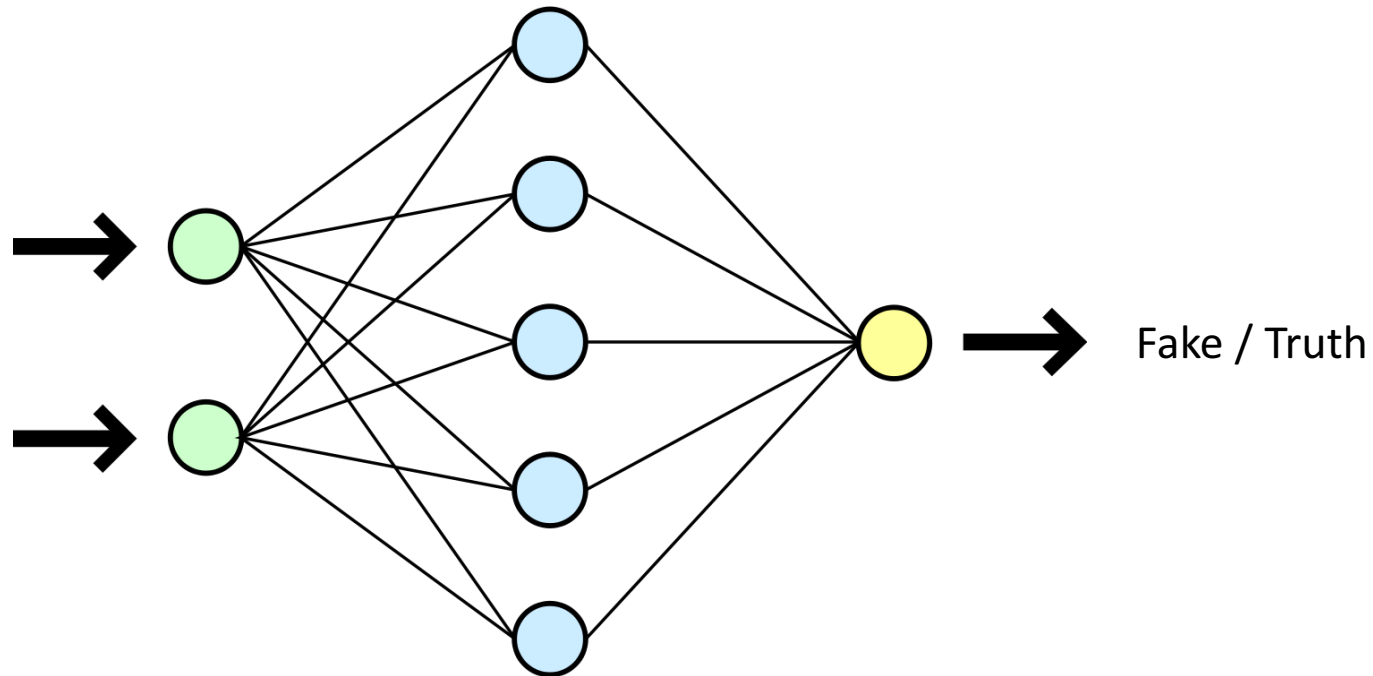
IBM Watson Tone Analyzer



Практична реалізація

Тренування нейронної мережі

- Agreeableness
- Analytical
- Anger
- Confident
- Conscientiousness
- Disgust
- Emotional Range
- Extraversion
- Fear
- Joy
- Openness
- Sadness
- Tentative



Практична реалізація

FakeCheck

Попередження: FakeCheck є науково-дослідницьким проектом на ранній стадії розробки, тому результати його роботи не є абсолютними і безсумнівними.
Завжди перевіряйте джерело інформації самостійно для прийняття остаточного рішення щодо його надійності.

Результати:

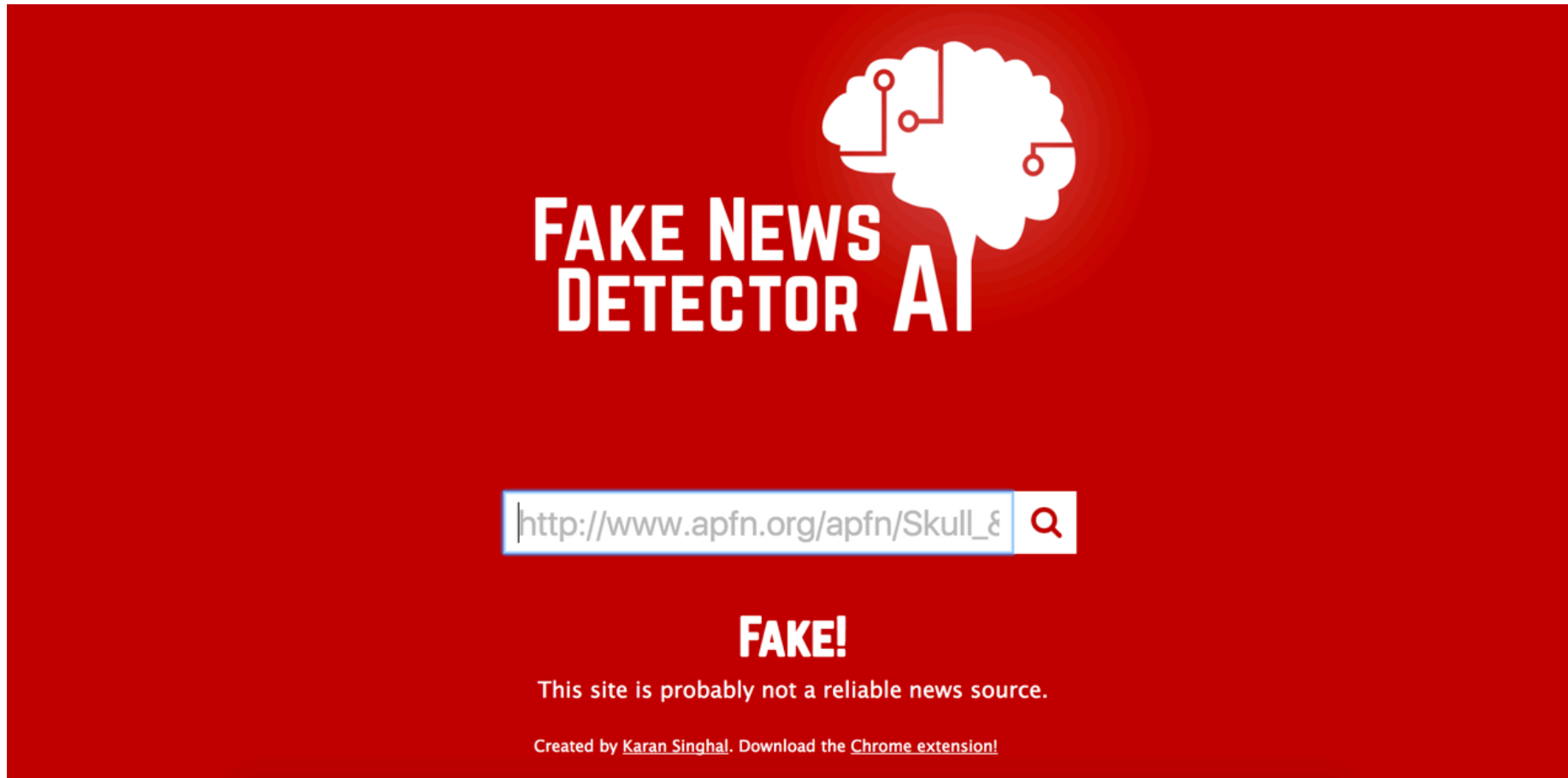
URL: http://www.apfn.org/apfn/Skull_&_Bones.htm

Це джерло скоріш за все є підробкою!

За результатами аналізу, це джерело є підробкою по причині:

Тест	Результат
Tone Analysis:	✘ Не пройдено

Порівняння з іншими рішеннями



**FAKE NEWS
DETECTOR AI**

http://www.apfn.org/apfn/Skull_&

FAKE!

This site is probably not a reliable news source.

Created by [Karan Singhal](#). Download the [Chrome extension!](#)

Результати

- Проведено аналіз існуючих систем перевірки правдоподібності інформації в мережі Інтернет
- Запропоновано метод, який відрізняється від багатьох існуючих систем тим, що виконує оцінку правдоподібності текстової в автоматичному режимі
- Новизна запропонованого методу полягає у використанні параметрів тону тексту в якості вхідних даних до нейронної мережі
- Запропонований метод показав свою дієздатність, отримані результати співпадають з результатами вже існуючих систем автоматичної перевірки Інтернет-сайтів на правдоподібність
- За даними експерименту встановлено зв'язок між рівнем правдоподібності текстової інформації та тоном, в якому вона була написана.

Майбутні дослідження

В майбутньому планується розширити навчальну вибірку і створити велику базу даних статей з їх рівнем правдоподібності.

Ще один напрямок розвитку - реалізація OCR з метою отримання додаткової інформації з фотографій в статтях, які можуть бути проаналізовані аналогічним чином до тексту статті. Аналіз зв'язків із зовнішніми джерелами інформації для того.

Для покращення результатів, необхідно визначити загальну ідею, яка не обов'язково безпосередньо згадується в тексті і використати це в якості додаткового набору вхідних даних для нейронної мережі використовуючи потужності IBM Watson Alchemy Language.

Дякую за увагу
