

## АРХІТЕКТУРА ГОЛОСОВОГО АНАЛІТИКА ПІД ЧАС ВИВЧЕННЯ ІНОЗЕМНОЇ МОВИ

Жаркіх С. Є.

Науковий керівник – ст. викл. каф. ШІ Гриньова О. Є.

Харківський національний університет радіоелектроніки, каф. ШІ

м. Харків, Україна

e-mail: [sofiia.zharkikh@nure.ua](mailto:sofiia.zharkikh@nure.ua)

Virtual Voice Analyst in learning English is an innovative technology created to transform the perception of studying process. Based on an Automatic Speech Recognition technology it helps to define both your knowledge and pronunciation level along with giving recommendations about further steps.

Англійська мова вже давно стала мовою комунікації, об'єднуючи представників різних культур та націй, а також своєрідним ключем до безлічі можливостей у різних сферах діяльності. Багато інновацій, технологічних рішень та винаходів відбуваються саме в англійськомовному середовищі. Знання англійської допомагає фахівцям, бути в курсі останніх тенденцій, взаємодіяти з колегами та фахівцями з усього світу, а також використовувати ресурси та документацію, що часто представлені саме англійською мовою, як на паперових носіях, так і в інтернеті.

Майже кожен другий опитаний українець (49%) не має жодних навичок англійської мови, тобто не може навіть прочитати і зрозуміти короткий текст, і тільки 23% можуть читати, писати і спілкуватися цією мовою на побутовому і навіть на професійному рівнях [1]. Брак кваліфікованих фахівців в ІТ-сфері із знаннями англійської мови є дуже актуальною проблемою. Зростання попиту на ІТ-працівників, які володіють не лише технічними навичками, але і англійською мовою на високому рівні, призводить до виникнення прогалини між вимогами ринку праці та рівнем підготовки фахівців. Аналогічно це стосується й інших сфер діяльності, адже в рамках міжнародної інтеграції та глобального партнерства простежується виникнення великої кількості дипломатичних, правових, економічних та технічних колаборацій. Відсутність достатнього рівня знань англійської мови обмежує можливості як сприйняття інформації, так і активної участі в обговореннях, переговорах чи обміні досвідом.

Зростаюча глобалізація та взаємодія країн роблять знання іноземних мов необхідністю. У сучасному світі інформаційних технологій (ІТ), де співпраця та обмін ідеями – це ключові складові успіху, володіння англійською мовою стає стратегічною конкурентною перевагою для фахівців різних галузей.

Вивчення іноземної мови – це процес, який, перш за все, має бути комфортним. На комфорт під час занять як з приватними викладачами, так і з представниками навчальних закладів впливає багато факторів: потреба у

гнучкому графіку, індивідуальний підхід чи психологічні особливості за типом особистості (інтроверти). Так традиційні заняття, де вчителем виступає жива людина, можуть виявитися обмеженими в порівнянні з використанням віртуального наставника, що створений на основі інноваційних технологій в галузі штучного інтелекту (AI).

Першим кроком людини, яка вирішила стати на шлях вивчення іноземної мови, зазвичай є аналіз вже набутих знань. На даному етапі в нагоді стає програмний додаток, який визначає рівень володіння мовою на основі записаного аудіо. Потенційний користувач мовного курсу робить запис своєї промови англійською на будь-яку з передбачених сервісом тем, при цьому використовуючи якомога більше різних структур в силу своїх максимальних знань. В свою чергу додаток «Голосовий аналітик», розроблений на основі моделей штучного інтелекту, розпізнає та верифікує як правильність вимови, порівнюючи зі зразком із своєї бази знань, так і робить морфологічний й синтаксичний аналіз записаного аудіо. Користувач отримує результат аналізу у вигляді текстового повідомлення у застосунку.

Технічна складова роботи голосового аналітика заснована на Automatic Speech Recognition (ASR) – технології машинного навчання для перетворення людської мови на читабельний текст. ASR є частиною галузі обробки природної мови Natural Language Processing (NLP), який в свою чергу вважається переважним напрямком під час процесингу даних у вигляді тексту [2].

Сам механізм аналізу мови полягає у розробці та навчанні двох моделей, на вході у яких буде записане користувачем аудіо. Перша передбачає використання правильного порядку слів, перевірку на коректність граматичних часів та на словниковий запас, а друга розбирає правильність вимовлених звуків, наголосу та інтонації (наприклад, як в запитальних реченнях).

Робота над першою моделлю починається в першому блоці, де відбувається парсинг фонем, тобто початковий аудіофайл, який складається з речень, ділиться на звукові слова, паузи, а також довгі паузи, що сигналізують про кінець речення. Оброблені та розділені звукові дані проходять етап з моделлю Voice-To-Text, що базується на технології ASR, і в результаті перетворюються на текстові слова, що є складовими речень. На етапі верифікації тексту за допомогою морфологічного та синтаксичного аналізатора створюється перший формальний звіт, який виводить результати та власне помилки, допущені користувачем.

Друга модель полягає в тому, що зразок звукового слова після етапу парсингу ставиться у відповідність текстовому слову, розпізаному першою моделлю, та порівнюється зі зразками із бази знань, що додатково містяться в системі. Як результат, формується другий формальний звіт, що фіксує та виводить помилки, які в подальшому будуть використовуватися під час генерації завдань за участю проблемних слів (рис.1).

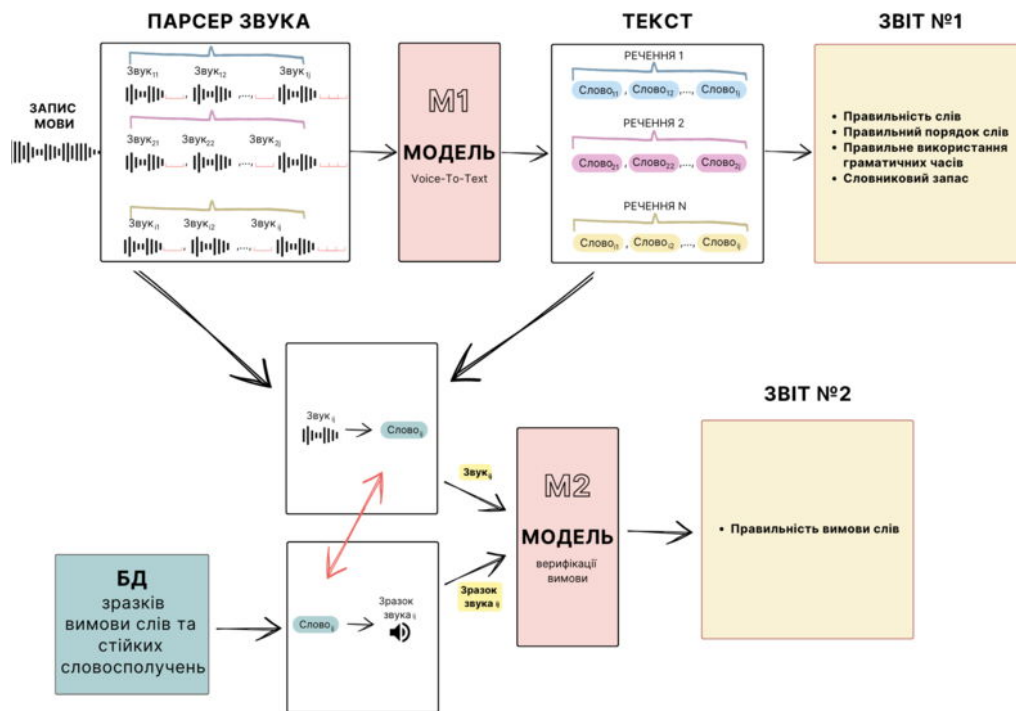


Рисунок 1 – Прототип архітектури «Голосового аналітика»

Розробка та навчання моделей, що оцінюють рівень знань англійської мови на основі обробленого аудіозапису полягає у використанні глибоких нейронних мереж, таких як Recurrent Neural Network (RNN) та Long Short-Term Memory (LSTM). LSTM є різновидом RNN і має механізм довгострокової пам'яті та призначений для управління проблемою загасання градієнтів у довгих послідовностях даних. Це дає змогу ефективніше працювати з довгими текстами, зберігаючи важливі контекстуальні залежності [3].

Таким чином, віртуальний викладач англійської мови, завдяки своїй гнучкості, передовим технологіям та індивідуальному підходу, започатковує новий стандарт у сфері навчання.

Список використаних джерел:

1. Оцінка рівня володіння іноземними мовами дорослого населення України. Звіт кількісного соціального дослідження. URL: <https://pidgorodne.otg.dp.gov.ua/storage/app/sites/92/uploaded-files/kmis-osinka.pdf> (дата звернення: 25.02.2024).

2. Lai J., Karat C.-M., Yankelovich, N. Conversational speech: interfaces and technologies Ed. By A.Sears, J. A Jacko // Human-computer interaction handbook: Fundamentals, evolving: technologies, and emerging applications. 2-nd ed.. P. 381-391. New York, 2008.

3. Jason Brownlee Long Short-Term Memory Networks With Python: Develop Sequence Prediction Models with Deep Learning. 2017. 229 p.