# 7.3. Обробка природної мови

Обробка природної мови (NLP - Natural Language Processing) - область, що знаходиться на перетині інформатики, штучного інтелекту та лінгвістики. З розвитком голосових інтерфейсів і діалогових систем, NLP стала однією з найважливіших технологій штучного інтелекту. Повне розуміння та відтворення змісту мовного фрагменту - надзвичайно складне завдання, оскільки людська мова має особливості:

* Це специфічна система усвідомленої передачі змісту сказаного або написаного.
* Це символьна система, для якої притаманна адаптивність, надійність та простота для опанування.
* Символи мови використовуються для спілкування по кількох каналах: звук, жести, текст, зображення і подібне.
* Природній голос не вимагає спеціальних знань і навичок, забезпечує всіх людей рівними можливостями.

Технології обробки природної мови давно вирішують вузькі завдання: пошукові системи обробляють запити з врахуванням семантики мови, сучасні перекладачі пристойно відображають загальний зміст фрази на інших мовах, голосові асистенти роблять менше помилок, а спеціальні архітектури нейронних мереж дозволяють витягувати актуальну інформацію з довільних текстів.

Основним завданням в NLP на сьогодні постає створення універсальних мовних моделей і архітектур, які будуть оперувати з текстом і голосом за допомогою однієї системи “End-To-End”. Системи, які будуть «розуміти» інформацію і зможуть взаємодіяти з користувачем подібно до людини з певним багажем знань.

#### Основні напрямки досліджень NLP

* Мовні технологій:
	+ Розпізнавання мови (STT - Speech-to-Text).
	+ Синтез мови (TTS - Text-to-Speech).
	+ Голосова біометрія (Voice Biometry).
* Розуміння природної мови (NLU – Natural Language Understanding).
	+ Голосовий інтерфейс (Voice User Interface).
	+ Діалогові системи, голосові помічники, чат-боти, системи «Питання-Відповідь».
	+ Машинний переклад (Machine Translation).
* Інші сфери взаємодії «людина-машина».

## Мовні технології

Мовні технології складаються з кількох складних напрямів.

* **Розпізнавання мови.** Перетворення мовного повідомлення у текст, бажано з максимальною точністю і якістю.
* **Синтез мови.** Завдання, зворотне до розпізнавання – озвучення тексту голосом так, щоб було зрозуміло для людини.
* **Голосова біометрія.** Визначення за голосом багатьох важливих характеристик людини: стать, вік, емоційний стан та ідентифікація особи.

На базі мовних технологій реалізуються алгоритми розуміння природної мови та складні системи: голосові інтерфейси, діалогові системи, інтелектуальні помічники, машинні перекладачі тощо.

## Розпізнавання мови

Розпізнавання мови - це процес перетворення мовного сигналу в цифрову інформацію. Це є одним з найскладніших технічних завдань, що відноситься до систем штучного інтелекту. Машинне перетворення мови в текст і голосової відповіді увійшли в повсякденне життя і поставлені на комерційний потік.

Стрибок у розвитку мовних технологій стався завдяки зменшенню вартості обчислювальних ресурсів та використанню нейронних мереж для вирішення багатьох завдань. У розпізнаванні мови використовуються останні досягнення в області машинного навчання, що сприяють високій точності розпізнавання.

#### Основні сфери застосування машинного розпізнавання мови:

* **Ідентифікація особи.** Використовується великими банками для підтвердження особи користувача за голосовим відбитком, для голосового підпису, а також в системах безпеки.
* **Аналітика дзвінків і переговорів.** Призначена для оцінки відгуків клієнтів, підвищення якості роботи операторів, виявлення трендів при зверненнях до служби підтримки або відділів з продажу.
* **Голосове управління.** Застосовується в багатьох сферах людської діяльності: в Інтернеті для пошуку інформації, в побуті для керування домашніми приладами, в системах «розумний будинок», навігаторах, управління автотранспортом, в мобільних додатках для виконання певних завдань.

На сьогодні сформувалося багато методів розпізнавання мови, у загальному їх можна розподілити на наступні типи:

* **Розпізнавання голосових міток та окремих команд.** Побудовано на роздільному проголошенні і подальшому розпізнаванні слова або словосполучення з невеликого тематичного словника. Використовується у відносно простих системах, призначених для виконання заздалегідь записаних голосових команд.
* **Пошук ключових слів.** Розпізнавання окремих ділянок мови, які містять задані слова чи словосполучення. Широко використовується контекст, підлаштування за рахунок додаткових даних, наприклад, контекст фрази або використання тематичного словника для даної предметної області. Застосовується в пошукових системах, а також в системах моніторингу мовлення.
* **Розпізнавання за граматикою.** Розпізнавання фраз, що відповідають певним заданим правилам (граматиці). Широко застосовується в системах голосового самообслуговування: користувача можуть попросити вимовити дату, будь-які номери, прізвища, адреси, підтвердити будь-яку дію словами «так» або «ні».
* **Розпізнавання лексичних елементів мови.** Найскладніша технологія: вона переводить мову в текст, не обмежуючись при цьому заданою граматикою. При його реалізації з потоку мови виділяються окремі звуки і з них складаються слова. Ситуацію ускладнює те, що одна і та ж фраза, вимовлена ​​різними людьми в різних умовах, буде звучати по-різному і надавати несхожі результати. Саме цей підхід і використовується в сучасних системах розпізнавання мови, що надають високу якість та достовірність розпізнавання.

## Синтез мови

Синтез мови (перетворення тексту в голос) на сьогоднішній день застосовується в самих різних областях. Це і голосові асистенти, і IVR-системи (Interactive Voice Response, інтерактивне голосове меню), і розумні будинки, і багато інших сервісів. Написаний текст повинен вимовлятися природньо правдоподібно до людської промови.

В напрямку синтезу голосу пройдено довгий шлях і застосовано багато методів, першим з яких був примітивний підхід склеювання окремих звуків в слова

### Метод «Unit-Selection»

Якщо набір текстів, які треба озвучити, є відносно невеликим і там зустрічаються повторні висловлювання - як, наприклад, в оголошеннях про відправлення і прибуття поїздів на вокзалі, - достатньо запросити диктора, записати в студії потрібні слова і фрази, а потім збирати з них повідомлення.

Диктор в студії наговорює багато текстів, всі промовлені слова розділяються на елементарні елементи (наприклад, на фонеми). Складається велика база даних: як звучить кожна фонема в залежності від того слова, в якому вона знаходиться. І коли виникає завдання озвучити якийсь текст, з цієї бази збираються і склеюються звуки.

На коротких текстах це працює пристойно: стандартні фрази: «Вас вітає компанія така-то» або «Сьогодні погода така-то». Але це не синтез мови, а просто відтворені частинки реальної мови. Такий підхід не працює з довільними текстами, якщо таким синтезом читати щось довге, наприклад, рекламне оголошення на два речення. Тоді у синтезі починає спотворюватися інтонація або голос, оскільки немає алгоритму, який відповідає за правильне об’єднання інтонації на стиках змістовних фрагментів фраз чи речень. Синтез довгого аудиотексту (наприклад, озвучити книгу) швидко починає дратувати, оскільки інтонація є монотонною і людина втомлюється це чути.

На сьогодні для синтезу мови використовують машинне навчання та нейронні мережі, що надають пристойну якість і наближений до людського синтезований голос.

### Статистичний (параметричний) метод

Статистичний (параметричний) синтез мови аналогічний до розпізнавання, тільки відбувається в зворотному напрямку. Є акустична модель, яка навчається на даних - записи мови з текстовими транскрипціями. Для збору даних диктор читає текст, але в значно меншому обсязі. Наприклад, відбувається запис 600 речень, що промовлені непрофесійним диктором за один день в студії, здійснюється синтез і отримання голосу.

Статистичний метод дозволяє швидко розширювати систему. У моделі Unit Selection для додавання нового голосу потрібно створити окрему мовну базу. Диктор повинен записати багато годин мовлення, при цьому бездоганно витримуючи однакову інтонацію. У статистичному методі для створення нового голосу достатньо записати хоча б дві години мовлення - приблизно тисяча вісімсот спеціальних, фонетично-збалансованих речень.

Такий підхід надає переваги. По-перше, швидкість відтворення різних голосів. По-друге, це завжди чітка і однакова інтонація на будь-якій довжині тексту. Голос в такому синтезі виходить з роботизованим забарвленням, хоча з розбірливістю та інтонаціями все добре. Завданням є знайти алгоритми, які допоможуть зробити синтезований голос максимально натуральним.

### Застосування нейронних мереж

На сьогодні в області синтезу мови, як і в багато інших областях широко застосовується машинне навчання. З'ясувалося, що цілий ряд компонентів всієї системи можна замінити на нейронні мережі, що дозволить не просто наблизитися за якістю до існуючих алгоритмів, а навіть значно їх перевершити.

Популярними архітектурами для синтезу на сьогодні є рекурентні мережі. На них вирішують всі завдання, пов'язані з послідовностями - голосові, текстові або музичні.

Сучасні голосові помічники навчають говорити не буквами, а фразами. Відразу змусити нейромережу цілком видавати фрази не вийде, адже тоді їй треба буде запам'ятати всі фрази в мові, цей розмір буде велетенським. Тут на допомогу приходить те, що текст, мова або музика - це послідовності. Кожне слово або звук - це самостійна одиниця, але яка залежить від попередніх.

Досить легко навчити мережу вимовляти окремі слова або букви. Береться множина аудіо файлів, розмічених на слова і навчається за вхідним словом видавати послідовність сигналів, схожих на його вимову. Здійснюється порівняння з оригіналом від диктора і намагання максимально наблизитися до ідеалу.

Озвучені тексти для навчання почали брати звідки завгодно. На прикладі багатьох відеозаписів з виступами Обами непогано навчили нейромережу розмовляти його голосом, що довело спроможності сучасних синтезаторів до імітації голосів.

## Голосовий інтерфейс

**Голосові інтерфейси (VUI - Voice-User Interface)** - це програмний продукт, який за допомогою голосової або мовної платформи дозволяє взаємодіяти користувачу і комп'ютеру, запускаючи автоматизовані процеси. Завдання таких інтерфейсів - розпізнати голос людини і виконати певне завдання.

Голосові інтерфейси зручні, коли вводити текст складно або незручно. Наприклад, під час водіння автомобіля користувач може проговорити свій запит, продиктувати потрібну адресу, перевірити завантаженість маршруту в додатку навігатора. Або якщо користувач виконує дуже багато завдань і не може сконцентруватися на одному.

#### Складові голосового інтерфейсу

1. **Голосове введення.** Запити вимовляються голосом, а не вводяться за допомогою клавіатури чи графічних елементів екранного інтерфейсу.
2. **Природна мова.** Користувачі не повинні обмежуватися використанням певного, оптимізованого для комп'ютера словника або синтаксису і можуть структурувати речення будь-якими способами, подібно до розмови з людиною.
3. **Голосове виведення.** Результат виконання вимовляється голосом, а не виводиться текстом на екрані.
4. **Інтелектуальна інтерпретація.** Для справжнього розуміння запитів користувача голосовий інтерфейс повинен використовувати додаткову інформацію, таку як контекст використання або дії, які користувач здійснював раніше.
5. **Сприяння.** Голосовий інтерфейс вчиняє дії необхідні для виконання завдання користувача, які користувач не запитував.

Голосовий інтерфейс зі всіма п'ятьма властивостями надає взаємодію з двома суттєвими перевагами:

* Можливість формулювати мету природньою мовою. Зникає необхідність у вивченні інтерфейсу і натисканні кнопок.
* Можливість передбачати цілі користувача, пропонувати додаткові дії на основі контекстної інформації або попередніх кроків.

### Голосове управління

Голосове управління - це вербальний процес спілкування з машиною. В хорошому голосовому інтерфейсі це спілкування має бути таким же природним, як і з людиною. У проектуванні таких систем значно більше психології та розуміння особливостей людського мислення. Розробники при проектуванні мають дотримуватися наступних принципів:

* Довіра
* Невидимий інтерфейс
* Людяність
* Індивідуальність

**Siri** - проект компанії, ідеологія якої - все повинно просто працювати. І все дійсно працює, якщо користувач вгадує з граматикою і словником. Якщо не вгадує, то система, без жодних пояснень на те, що пішло не так і як скорегувати поведінку, просто перестає працювати.

При цьому великого акценту зроблено на індивідуальності. Якість голосу, жарти, смішні коментарі при виконанні звичайних завдань вражають. Користувач розслабляється і намагається взаємодіяти з Siri як з особистістю. Але коли система починає реагувати не так, як він очікує, сприйняття різко зменшується. Користувач думає, що його дії не схвалюються або над ним просто сміються, і це значно гірше, ніж якби він спочатку сприймав її як машину.

В Google порахували більш безпечним не намагатися імітувати індивідуальність, а показувати, що перед користувачем просто високотехнологічний програмний продукт, в якого навіть немає імені (OK, Google).

### Розуміння природної мови

На етапі розуміння природної мови пристрій має справу вже не з самою мовою, а з текстом, в який вона була перетворена. Системи розпізнавання мови можна використовувати не тільки в голосових інтерфейсах, але і для створення ботів, здатних спілкуватися або відповідати на листи.

Звичайно, самі по собі слова для комп'ютера нічого не значать. Подібно до людини, яка бачить перед собою текст, написаний невідомою мовою, не може уявити, про що йдеться, якщо тільки не побачить знайомі слова. Приблизно таке ж завдання виконує комп'ютер - за знайомими словами і формулюваннями зрозуміти вимовлене.

Щоб зрозуміти, чого хоче користувач, і активувати потрібну функцію, система має впізнати в мові ключові слова, наприклад, «виклич таксі». Цим словам його навчив розробник: вони занесені в програму в числі інших формулювань виклику таксі: «замов таксі», «виклич машину», [...] «хочу поїхати додому». Природно, всіх варіантів розробник врахувати не може. Тому, зібравши тестову версію додатку, пропонується покористуватися ним багатьом людям - так з'ясовується, як ще люди можуть замовляти таксі.

Всі нові формулювання записуються в конфігураційний файл програми, а потім розмічаються: в них виділяються слова-маркери і характерні синтаксичні схеми. Аналізуючи розмітку, програма сама вчиться розуміти навіть такі формулювання, яких немає в списку, але які схожі на відомі їй конструкції. Наприклад, якщо в списку є варіанти «виклич машину» і «замов таксі», то і формулювання «замов машину» програма теж зрозуміє.

### Модуль голосової активації

Друга ключова складова голосового інтерфейсу - це система активації голосом, яка запускає потрібну дію у відповідь на ключову фразу. Розробники Google для розпізнавання коронної фрази «Окей, Google» використовують глибоку нейронну мережу, яка надає високу якість. Але система активації при цьому обмежена єдиною командою, і для навчання потрібна величезна кількість даних. Модель для розпізнавання всім знайомої фрази навчалася на прикладі більш ніж 40 тисяч голосів користувачів, які зверталися до своїх смартфонів з Google Now.

В інших голосових інтерфейсах застосовують різні підходи. Наприклад, користувач самостійно обирає ключову фразу з наявної бібліотеки або придумує свою. Розпізнавання команди повинно відбуватися на самому пристрої, без звернення до сервера (рис.1).



Рис.1. Розпізнавання асистентом ключової фрази

Для системи активації потрібен невеликий словник, адже їй достатньо розуміти кілька ключових фраз, а всю іншу промову можна просто ігнорувати. Тому, мовна модель активації є компактною. Якщо повноцінна модель розпізнавання може мати десятки мільйонів станів, що описують як чітку мову так і тишу, сторонній шум, фонові розмови та займати гігабайти, то для голосової активації вона обмежена сотнями станів і вміщується в кілька десятків кілобайт.

Модель для розпізнавання ключової фрази має враховувати особливості вимовляння навіть й при невеликому словнику. Команда повинна бути відносно довгою (бажано - більше одного слова) і нечасто зустрічатися в повсякденній мові, щоб виключити помилкові спрацьовування. «Будь ласка» погано підійде для голосової активації, а «слухай мою команду» - цілком.

В систему вбудовано детектор голосової активності, який стежить за появою ключової фрази у вхідному звуковому потоці людського голосу. Решта звуків ігноруються, тому в фоновому режимі енергоспоживання модуля активації обмежено тільки мікрофоном.

### Виділення змістовних об'єктів

Слова, які вимовляє людина, важливо не лише перевести в літери, а й наповнити змістом. Ця технологія починає працювати вже після того, як вимовлені слова розпізнано. Повний стек мовних технологій містить модуль виділення змістовних об'єктів в природному мовленні, і на виході отримується не просто розпізнаний, а вже розмічений текст.

Як змістовні об’єкти виділяються дати, час, адреси, імена, прізвища, географічні назви. У гібридній системі поєднуються контекстно-вільні граматики, словники ключових слів і статистичні дані пошуку і різних сервісів, а також алгоритми машинного навчання. Наприклад, у фразі «поїхали на головний залізничний вокзал» слово «вокзал» допомагає системі визначити контекст, після чого в базі даних знаходиться відповідний об'єкт (рис.2).



Рис.2. Виділення змістовних об'єктів

На вхід системи подається список різних типів об'єктів і приклади фраз з живої мови, що їх описують. Далі з цих прикладів формуються патерни, в яких враховуються початкова форма, корінь, морфологічні варіанти слів. Наступним кроком надаються приклади вживання обраних об'єктів в різних поєднаннях, які допоможуть системі розуміти контекст. На основі цих прикладів будується модель, де виділені в репліці користувача об'єкти стають станами, що відстежуються.

## Загальні технологічні тренди NLP

#### End-To-End рішення задач NLP

Все більше рішень засновано на End-To-End принципі, наприклад нейромережна модель отримує на вході акустичний сигнал (звукові хвилі) і видає на виході акустичний сигнал, без проміжної фази тексту. Це істотно пришвидшує час виконання та якість, але одночасно погіршує «прозорість» і розуміння того, «що всередині».

#### Наближення якості розпізнавання і генерації мови до людської

Спостерігається суттєве покращення якості розпізнавання мови. Кількість помилок при розпізнаванні змішаної мови кількох людей, які говорять з різними акцентами в умовах зашумлености наближається до людського рівня. Відбувається аналіз звукових сцен з розпізнаванням статі і віку осіб, емоційного забарвлення розмови та оточуючий фон. Синтезована мова вже спроможна відтворювати голос будь-якої людини.

#### Багатомовність

З'являються багатомовні моделі перекладу, в тому числі, за рахунок застосування Transfer Learning і за рахунок використання окрім паралельних корпусів значно більше об'ємних монокорпусов. В результаті підвищується якість перекладу для малоресурсних мов (з відносно невеликими масивами навчальних вибірок).

Ручний переклад витісняється машинним завдяки більш глибокому машинному розумінні контексту і тематики документів. Очікується машинний синхронний усний переклад.

#### Розуміння змісту текстів

З'являються додатки, що засновані на розумінні змісту з врахуванням контексту: різного роду діалогові та довідкові сервіси, здатні розуміти контекст діалогу, розумно відповідати на запитання користувачів і скеровувати хід діалогу в потрібному напрямку. Більш глибоке машинне розуміння мови виводить на новий рівень автоматичну обробку текстових потоків в інтернеті і в соціальних мережах: збір та компіляцію фактів, їх аналіз на несуперечливість і достовірність.

#### Генерація текстів

Нейромережні End-To-End рішення повсюдно заміняють класичні методики. Застосування нейромереж надає можливість створювати досить довгі статті на довільні теми в заданій області з керованим вмістом. Очікують на нейромережні моделі, що зможуть генерувати тексти не гірше за людину, і тоді автоматичний контент затопить світ.

#### Платформи і крос-платформність

З'являється багато платформ для складання додатків на базі голосових інтерфейсів. Хмарні платформи покращують показники щодо часу відгуку, навантажень і безпеки. Прогнозується зростання інвестицій не в окремі діалогові сервіси (чат-бот), а в багатофункціональні платформи і крос-платформні рішення, завдяки яким голосовий асистент зможе однаково працювати на різних пристроях.

#### Технології для малої кількості даних

Зростатиме значення методів машинного навчання, які ефективно працюють в умовах малої кількості сирих даних: Transfer Learning, Knowledge Transfer. В таких застосуваннях також очікується більш широке використання GAN (Generative Adversarial Networks) для генерації даних для навчання моделей.

#### Архітектури з меншими вимогами до обчислювальних ресурсів

У міру переходу нейромережних моделей зі стін лабораторій в комерційні дата-центри підвищуються вимоги до їх енергоефективності. Очікується поява нових, більш ефективних обчислювальних архітектур. Наприклад, розріджені мережі, що поєднують кращі якості розподілених і символьних обчислень, моделі, складність яких адаптується до кількості навчальних даних.

#### Повсюдне впровадження голосових інтерфейсів

Розвиток технологій Speech-to-Text є кроком для спрощення офісних завдань (наприклад, планування часу керівника, пошуку документів, обробки конфіденційної інформації). З підвищенням точності розпізнавання, глибини розуміння і якості синтезу мови голосові інтерфейси можуть бути інтегровані практично у всі пристрої: діалогові системи в розумному будинку, автомобілі, побутову техніку, боти-помічники.

#### Вибухове зростання голосових роботів

Зростає кількість інтелектуальних асистентів в різних галузях бізнесу, в тому числі в комерційних сервісах банків, торгових мереж, телекомунікаційних та інших компаній, які активно взаємодіють з клієнтами. Все вербальне спілкування з масовою аудиторією в найбільш популярних сервісах будуть вести роботи. Роботи навчаються точно розпізнавати емоції, використовуючи в тому числі мультимодальну оцінку емоцій і здатні самі використовувати емоційну складову в розмові.

#### Пошук інформації на природній мові

Зростає попит на інтелектуальний пошук з можливістю робити запити природньою мовою. Все більше організацій хочуть швидко знаходити неструктуровані дані у всіх внутрішніх джерелах, автоматично визначати їх зміст і виділяти значущі факти в спеціалізованих юридичних або фінансових текстах. За рахунок розвитку глибоких моделей витягування фактів з текстів і автоматичне реферування їх змісту буде істотно підвищуватися якість пошуку інформації.

## Висновки

NLP (Natural Language Processing, обробка природної мови) - це область обчислень, мета якої - допомогти комп'ютерам зрозуміти людську або «природну» мову. На даний момент хвиля інтересу до NLP тільки зростає, і у найближчому майбутньому голосова взаємодія отримає більшого поширення майже у всіх сферах діяльності.

Хоча природні мови не здаються чимось дивовижним, комп'ютерам дуже важко правильно їх інтерпретувати і використовувати. Строгий, обмежений правилами формат електронних таблиць і баз даних ідеальний для машин, а випадкова природа людських мов, що залежить від контексту і не завжди пов'язана певними правилами спантеличує штучний інтелект.

Звісно, NLP ще далекий від того, щоб змінити світ професій прямо зараз, але цей напрямок існує вже близько 30 років і роботи з його вдосконалення ведуться постійно. Експерти вважають, що наступний прорив у розвитку NLP буде колосальний, зумовить перехід від структурованих (бази даних) до неструктурованих (текст) даних і значно поліпшить здатність машин «розуміти» людей в звичайній розмові. Пристрої, що здатні розпізнавати голос і генерувати його, стрімко дешевшають і вдосконалюються.

## Контрольні питання

1. Які особливості має людська мова?
2. Назвати основні напрямки досліджень в NLP. Які напрямки формують мовні технології?
3. Назвати основні сфери застосування машинного розпізнавання мови.
4. Перелічити поширені методи машинного розпізнавання мови.
5. Які існують підходи до синтезу мови.
6. Назвати основні етапи для статистичного методу синтезу мови
7. Яку користь привносить застосування нейронних мереж для розпізнавання та синтезу мови.
8. Перелічити функції, що виконує голосова біометрія.
9. Назвати основні складові голосових інтерфейсів.
10. Перелічити технологічні тренди розвитку NLP.

## Використані джерела

1. Історія обробки природньої мови <https://habr.com/ru/post/481228/>
2. NLP-технології розпізнавання людського мовлення <https://evergreens.com.ua/ua/articles/natural-language-processing.html>
3. Обробка природньої мови, розпізнавання та синтез мовлення <https://singapore-academy.org/libcdo/100.pdf>
4. Мовні технології [https://www.tadviser.ru/index.php/Речевые\_технологии](https://www.tadviser.ru/index.php/%D0%A0%D0%B5%D1%87%D0%B5%D0%B2%D1%8B%D0%B5_%D1%82%D0%B5%D1%85%D0%BD%D0%BE%D0%BB%D0%BE%D0%B3%D0%B8%D0%B8)
5. Мовні технології в бізнесі <https://hightech.fm/2020/05/21/chatbot-bussines>
6. Сучасний ринок мовних технологій <https://www.controlengrussia.com/innovatsii/rechevye-tehnologii/>
7. Розпізнавання мови <https://habr.com/ru/company/toshibarus/blog/490732/>
8. Машинне навчання це весело! Частина 6: Розпізнавання мови за допомогою глибокого навчання [https://medium.com/@ppleskov/машинное-обучение-это-весело-часть-6-aa3982f25d54](https://medium.com/%40ppleskov/%D0%BC%D0%B0%D1%88%D0%B8%D0%BD%D0%BD%D0%BE%D0%B5-%D0%BE%D0%B1%D1%83%D1%87%D0%B5%D0%BD%D0%B8%D0%B5-%D1%8D%D1%82%D0%BE-%D0%B2%D0%B5%D1%81%D0%B5%D0%BB%D0%BE-%D1%87%D0%B0%D1%81%D1%82%D1%8C-6-aa3982f25d54)
9. 8 головних досягнень в нейромережному NLP <https://sysblok.ru/nlp/8-glavnyh-proryvov-v-nejrosetevom-nlp/>
10. 5 методів обробки природної мови <https://neurohive.io/ru/osnovy-data-science/5-metodov-v-nlp-kotorye-izmenjat-obshhenie-v-budushhem/>