

5. Kim JO (Ryan), Jeong YS, Kim JH, Lee JW. Machine Learning-Based Cardiovascular Disease Prediction Model: A Cohort Study on the Korean National Health Insurance Service Health Screening Database. *Diagnostics*. 2021 May 25;11(6):943. <https://doi.org/10.3390/diagnostics14020144>

6. Whiting DR, Guariguata L, Weil C, Shaw J. IDF Diabetes Atlas: Global estimates of the prevalence of diabetes for 2011 and 2030. *Diabetes Research and Clinical Practice*. 2011 Dec;94(3):311–21. doi: 10.1016/j.diabres.2011.10.029.

7. Ledisi G. Kabari: Diagnosing Diabetes Using Artificial Neural Networks. February 2020 *European Journal of Engineering and Technology Research* 5(2):221-224. DOI:10.24018/ejers.2020.5.2.1774

## **ГЛИБОКІ МОВНІ МОДЕЛІ: ВІД ТЕХНОЛОГІЧНОГО ПРОРИВУ ДО ПРАКТИЧНИХ ЗАСТОСУВАНЬ**

*Літвін Олександр*

*Харківський національний автомобільно-дорожній університет, Харків*

Глибокі мовні моделі - це сукупність алгоритмів машинного навчання, які призначені для розуміння та генерації природної мови. Вони базуються на глибокому навчанні (deep learning) і використовуються для розв'язання різноманітних завдань у сфері обробки природної мови (NLP), таких як машинний переклад, синтез мови, розпізнавання мовлення тощо.

### *Огляд поняття глибоких мовних моделей*

Глибокі мовні моделі відрізняються від традиційних методів обробки мови своєю здатністю адаптуватися та вчитися на великих обсягах даних. Вони використовують глибокі нейронні мережі з численними шарами (відси назва "глибокі") для автоматичного вивчення складних залежностей у тексті.

### *Історія розвитку в цьому напрямку*

Історія розвитку глибоких мовних моделей пов'язана з виникненням та постійним вдосконаленням штучних нейронних мереж. Одним з ключових проривів стало введення архітектури Transformer, яка стала основою багатьох сучасних моделей. Починаючи з моделі BERT від Google до GPT (Generative Pre-trained Transformer) від OpenAI, глибокі мовні моделі стали домінуючими в галузі обробки природної мови.

### *Важливість глибоких мовних моделей у сучасному світі*

Глибокі мовні моделі відіграють ключову роль у сучасному світі завдяки своїм можливостям в розумінні та генерації мови. Вони використовуються у багатьох галузях, від інтелектуальних асистентів до платформ штучного інтелекту. Їх важливість полягає в здатності автоматизувати та поліпшувати обробку природної мови, що відкриває нові можливості для розвитку технологій та покращення користувацького досвіду.

## **Технологічний прогрес у глибоких мовних моделях**

### *Архітектура моделей: Transformer*

Transformer - це архітектура глибоких нейронних мереж, яка вперше була запропонована у статті "Attention is All You Need" в 2017 році. Ця архітектура замінила традиційні рекурентні та зворотні нейронні мережі у задачах обробки природної мови. Основними складовими Transformer є механізм самоуваги (self-attention) та позиційне кодування, які дозволяють моделі ефективно опрацьовувати послідовності будь-якої довжини.

### *Підходи до навчання: Self-attention mechanism*

Механізм самоуваги (self-attention) є ключовим елементом архітектури Transformer. Цей механізм дозволяє моделі приділяти різній увазі різним частинам вхідного тексту під час обробки. Він дозволяє моделі "фокусуватися" на важливих аспектах вхідного тексту, забезпечуючи кращу якість розуміння та генерації тексту.

### *Шляхи оптимізації та розвитку моделей*

Шляхи оптимізації та розвитку глибоких мовних моделей включають в себе використання більших та якісніших наборів даних для навчання, оптимізацію алгоритмів навчання та пошук нових архітектур моделей. Також важливою є розробка методів передбачення та генерації тексту, які були більш точними та ефективними.

## **Використання глибоких мовних моделей у різних сферах**

### *Обробка природної мови (Natural Language Processing, NLP)*

Глибокі мовні моделі виявляються надзвичайно корисними в сфері обробки природної мови. Вони використовуються для розпізнавання та класифікації тексту, визначення настрою, екстракції інформації, сумаризації тексту та багатьох інших завдань. Завдяки глибоким мовним моделям можливо автоматизувати багато процесів, що пов'язані з обробкою тексту, що робить їх незамінними в багатьох сферах, від клінічної медицини до фінансів та маркетингу.

### *Машинний переклад*

Глибокі мовні моделі перевершили традиційні методи машинного перекладу, забезпечуючи більш точні та природні переклади. Вони використовують контекстуальну інформацію та розуміння мови для генерації більш вірних перекладів, що робить їх незамінними для комунікації між мовами у різних галузях, від бізнесу до туризму.

### *Синтез мови*

Глибокі мовні моделі також застосовуються у синтезі мови, тобто у створенні тексту на основі вхідних даних. Вони можуть бути використані для автоматичного створення описів продуктів, генерації новинних статей, писання креативного контенту та багатьох інших завдань, які вимагають генерації тексту.

### *Генерація тексту*

Глибокі мовні моделі відіграють ключову роль у генерації тексту для різних цілей, від створення автоматичних відгуків у соціальних мережах до генерації творчого контенту у мистецтві та літературі. Вони можуть бути навчені на великих обсягах текстових даних та створювати текст, який максимально наближений до стилю та тематики вихідних даних.

### *Розпізнавання мови*

Глибокі мовні моделі використовуються для розпізнавання мовлення у різних ситуаціях, від голосових асистентів на смартфонах до систем автоматичного підпису. Вони дозволяють точно розпізнавати інтонацію, акцент та інші аспекти мовлення, що робить їх корисними для різних завдань у галузі комунікації та інтерфейсів користувача.

Глибокі мовні моделі мають широкий спектр застосувань у різних сферах, від обробки природної мови до машинного перекладу та генерації тексту. Їх потужність та гнучкість роблять їх незамінними у сучасному світі комунікації та обробки інформації.

### **Виклики та перспективи**

#### *Обмеження сучасних моделей*

Незважаючи на значний технологічний прогрес, сучасні глибокі мовні моделі мають свої обмеження. Одним з найбільших обмежень є нестабільність у роботі з малочисельними або неправильно підготовленими даними, що може призводити до неправильних результатів або викривлення виходу моделі. Також важливими є обмежені можливості мовних моделей у врахуванні контексту та семантики, що обмежує їхню точність у ряді завдань.

#### *Етичні питання використання глибоких мовних моделей*

Зростаюча потужність глибоких мовних моделей ставить перед нами ряд етичних питань. Наприклад, можливість створення фальшивих новин або маніпуляції громадською думкою за допомогою автоматичної генерації тексту. Також важливими є питання приватності та захисту персональних даних, оскільки глибокі мовні моделі можуть потенційно здійснювати аналіз та інтерпретацію особистої інформації.

#### *Можливі напрямки подальшого розвитку*

Незважаючи на виклики, перед глибокими мовними моделями відкриті широкі перспективи розвитку. Одним із напрямків є пошук нових архітектур моделей, що забезпечать кращу точність та гнучкість у різних завданнях. Також важливим є розробка методів адаптації моделей до конкретних доменів або мовних варіантів, що дозволить їх ефективніше використовувати у практичних застосуваннях. Подальший розвиток глибоких мовних моделей також пов'язаний з покращенням їхньої здатності до розуміння та генерації мови, що відкриває нові можливості для застосування в різних сферах, від освіти до медицини.

Хоча глибокі мовні моделі зустрічаються з викликами, вони також мають великий потенціал для подальшого розвитку та застосування в різних сферах життя. Розуміння цих викликів та перспектив може допомогти нам вдосконалити та використовувати ці моделі з максимальною користю для суспільства.

### **Приклади успішного застосування глибоких мовних моделей**

Кращі практики у використанні глибоких мовних моделей

Модель BERT від Google відзначається використанням у класифікації тексту з вражаючою точністю, зокрема у сферах аналізу соціальних медіа та медичного тексту.

Netflix успішно використовує моделі на основі Transformer для рекомендацій фільмів та серіалів, що дозволяє надавати персоналізовані рекомендації користувачам.

Кейси успіху у відомих компаніях або дослідницьких проектах

1. **OpenAI і модель GPT:** Модель GPT від OpenAI використовується для створення природних текстових відповідей та новинних статей з вражаючою автономністю.

2. **Microsoft і проект Turing-NLG:** Проект Turing-NLG від Microsoft створює AI-систему, здатну відповідати на питання користувачів у реальному часі.

3. **Facebook та модель BART:** Модель BART від Facebook використовується для машинного перекладу, генерації тексту та сумаризації, покращуючи допоміжні інтерфейси та інтерактивні системи.

### **Джерела літератури**

1. Васвані, А., Шазір, Н., Пармар, Н., та ін. (2017). "Увага - це все, що вам потрібно".
2. Девлін, Дж., Чанг, М. У., Лі, К., та Тоутанова, К. (2018). "BERT: Переднавчання глибоких двонапрямлених трансформаторів для розуміння мови".
3. Редфорд, А., Ву, Дж., Чайлд, Р., та ін. (2019). "Мовні моделі - це ненавчені багатозадачні вчені". Блог OpenAI, 1(8), 9.
4. Льюїс, М., Ліу, Ї., Гойяль, Н., та ін. (2020). "BART: Шумопідготовка передавання послідовностей-до-послідовності для генерації природних мов, перекладу та розуміння".
5. Браун, Т. Б., Манн, Б., Райдер, Н., та ін. (2020). "Мовні моделі - це навчені вчені".