ADVANCED NLP AI TOOLS
IMPORTANCE IN MACHINE TRANSLATION FROM UZBEK TO ENGLISH

Nazirova Elmira
DSc, Head of the Department of Multimedia Technologies,
Tashkent University of Information Technologies
named after Muhammad Al-Khwarizmi.
E-mail: elmira_nazirova@mail.ru

Usmonova Kamola
PhD of the Department of Multimedia, Technologies,
Tashkent University of Information Technologies
named after Muhammad Al-Khwarizmi.
E-mail: kamolausmonova1234@gmail.com

Abstraction: This article explores the substantial influence of artificial intelligence (AI) on digital technology, particularly in the field of machine translation. It investigates the strides and obstacles in AI-powered machine translation, focusing specifically on Uzbek and English. It investigates the exciting developments in language translation technology, pushing the boundaries beyond the capabilities of current AI tools and Natural Language Processing. Recent strides in NLP, particularly in the domain of neural machine translation, have propelled us towards more accurate and contextually relevant translations. The integration of advanced NLP techniques, such as attention mechanisms and transformer architecture, has elevated the quality of translations, pushing the boundaries of what AI tools can achieve. Exciting developments in multilingual NLP models, fueled by AI-driven research, hold promise for overcoming existing limitations and unlocking new frontiers in cross-linguistic communication.

Keywords: Artificial Intelligence, Machine Translation, Uzbek, Russian, English, Natural Language Processing (NLP), Translation Technology, Linguistic Resources, Corpora.

Аннотация: В этой статье исследуется значительное влияние искусственного интеллекта (ИИ) на цифровые технологии, особенно в области машинного перевода. Она исследует прогрессы и препятствия в машинном переводе, основанном на ИИ, с особым акцентом на узбекском и английском языках. Также рассматриваются интересные нововведения в технологии языкового перевода, выходящие за рамки возможностей существующих инструментов ИИ и обработки естественного языка. Недавние достижения в области обработки естественного языка, особенно в сфере нейронного машинного перевода, подвигают нас к более точным и контекстно релевантным переводам. Интеграция передовых техник обработки естественного языка, таких как механизмы внимания и архитектура трансформера, повысила качество переводов, расширяя возможности инструментов ИИ. Захватывающие нововведения в мульти язычных моделях обработки естественного языка, стимулируемые исследованиями, основанными на ИИ, обещают преодолеть существующие ограничения и открыть новые горизонты в межъязыковой коммуникации.

Ключевые слова: Искусственный Интеллект, Машинный Перевод, Узбекский, Русский, Английский, Обработка Естественного Языка (NLP), Технология Перевода, Лингвистические Ресурсы, Корпусы.

Introduction

These days machine translation is very important for breaking down barrier between language. As Snell-Hornby (2006) mentioned that in the 1990s English became prevalent language for interacting as a technological lingua franca and allowed machine translation to localizing content. This shift was driven by the growing demand for rapid communication and globalization, highlighting the importance of natural language processing (NLP) techniques in understanding language structures while bridging cultural differences. Moreover, advancement of the evolution of technology designed to automatically interpret text or speech across diverse languages finds representation in machine translation—a fusion of artificial intelligence (AI) and language translation. As McKinsey (2018) pointed out that by digital productivity is expected to rise for the next decade by 60% in economic and social benefits. Enhancement and usage of AI and machine learning, with particularly focus on NLP-driven machine translation. Yet, as Galati and Riediger (2017) stated that there appears some threat to jobs and population, thus machine translation will take over human role in the upcoming future. Despite these challenges, AI-powered translation tools offer undeniable benefits. By leveraging NLP algoritms, they can quickely process and translate large volumes of text with remarkable efficiency. According to Forcada (2017) Neural Machine Translation (NMT) is clearly out-lined as a "corpus-based machine translation's new breed" and it demand extensive analysis of huge corpora of SL as well as TL sentences. As innovation computational approach, namely neural network is introduced. In addition, it is necessary for translation from Uzbek into English. While AI-powered translation tools have made remarkable strides, they face hurdles when dealing with intricate languages like Uzbek and English. These complexities have spurred an exciting journey towards a newer, more refined breed of translation technology that goes beyond mere word conversion.
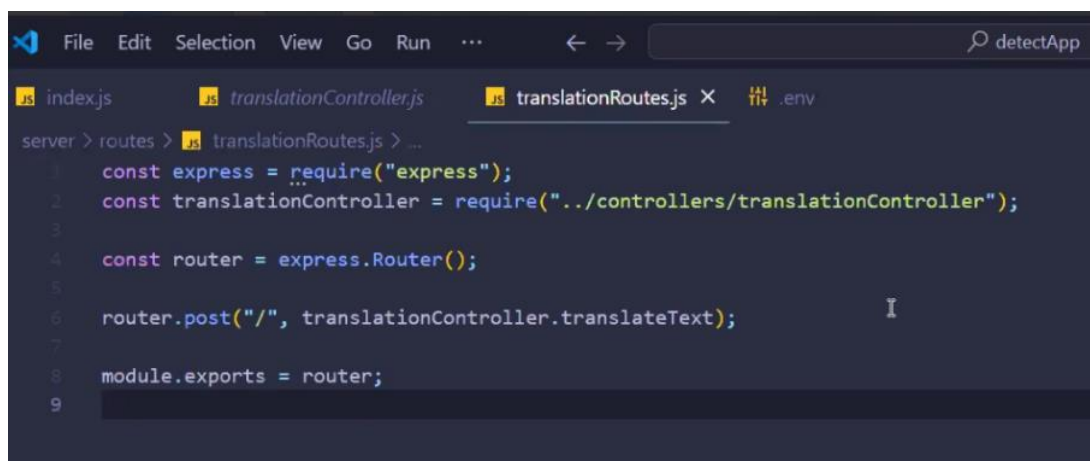
Challenges in Uzbek language translation

Unlike languages such as English, Uzbek lacks substantial linguistic assets like text corpora, annotated datasets, and lexicons, hindering the development of accurate NLP AI tools. Additionally, the complex morphology and syntax of Uzbek demand specific linguistic expertise and resources, further complicating the development of effective NLP AI tools. Except Uzbek, there is also some challenges in other language such as Russian translation. Languages rich in morphology, such as Russian, present complexities for machine translation systems. Vocabulary limitations fail to encompass Russian's intricate grammatical changes, affecting translation quality. Hence, it is necessary to use NLP models in order to train on large datasets can learn these patterns and provide contextually appropriate translation.

AI tools importance of Uzbek machine translation

As Munday (2012) pointed that translation has risen in sophistication with usage of technology as a tool for translation and it is explained scientific discipline that adduce both the theory and practice of translation. There are different tools for translating such as Google translator, Yandex and others. Yet, there are some challenges during translating using them. As Rudy and Christanta (2020) mentioned that through observation they find out that Google translate functions is not perfect translator in some cases as well as there are minor errors and diminished

precision. The reason of its inaccuracies, this translation tool indicates limited use. As a result, a mere alternation in translation meaning impact significantly the result's quality. Hence, the main barrier occurred when users use a translation tool, they do not get intended meaning because of conveying inaccurately. Moreover, In Uzbekistan, many people struggle translating some articles or different resources from Uzbek into English. The reason there is no trustworthy tool. This is because Uzbek and English structure very different with each other. Hence, there is some problems while translating using google or Yandex translator. Considering some translation tools such as Google translator, Yandex translator, UzbekVoice AI, MohirAI, Bing Microsoft translator, Taia, and others, we can notice inaccurate translation. For example, Google Translate struggles with conveying accurate meanings due to its inability to consider context. Complex texts and context-dependent nuances, uninterpretable without deep language knowledge, increase the likelihood of errors. Unlike professional translators who utilize various tools like glossaries, back translation, and reviewers to prevent literal and nonsensical translations, Google's direct translation often results in nonsensical outcomes. After translating, it is necessary to look through by someone and make context understandable. Hence, by harnessing the power of AI and NLP, researchers are pushing the boundaries of current capabilities, paving the way for more accurate, fluent, and culturally nuanced translation. The translation quality heavily relies on the specific language pair involved. Languages like Uzbek and English, more prevalent in Google's database, yield more accurate translations, whereas languages with limited database availability might suffer from inaccuracies. Grammatical errors are common in Google Translate as its system relies on language pair frequency rather than adhering to grammatical rules. This method leads to translations with significant grammatical flaws. Moreover, Google Translate lacks error correction systems. Users cannot report errors or proofread translations unless they are proficient in both the source and target languages. While creating machine translation tool we will get some words range from google and relook every word because there are some corrections needed. Yet, after deep analyses, we use our own algorithm in order to get more accurate output.

```
const express = require("express");
const translationController = require("../controllers/translationController");

const router = express.Router();

router.post("/", translationController.translateText);

module.exports = router;
```

```
File  Edit  Selection  View  Go  Run  ...        ←  →                                    ⌕ de

Js index.js  ×      Js translationRoutes.js    ⋔⋔ .env

server > Js index.js > ...
   1    const express = require("express");
   2    const translationRoutes = require("./routes/translationRoutes");
   3    const annyang = require("annyang");  4.3k (gzipped: 2k)
   4
   5    const app = express();
   6    const port = process.env.PORT || 3000;
   7
   8    app.use(express.json());
   9
  10    app.use("/translate", translationRoutes);
  11
  12    annyang.start();
  13
  14    annyang.addCommands({
  15      "translate *text to *targetLanguage": translateVoiceCommand,
  16    });
  17
  18    function translateVoiceCommand(text, targetLanguage) {
  19      console.log("Voice command detected:", { text, targetLanguage });
  20    }
  21
  22    app.listen(port, () => {
  23      console.log(`Server is running on http://localhost:${port}`);
  24    });
  25
```

Enter the new breed of translation tech

It is time to deep analysis the morphology and syntax both language and create tool, which make translation clearer. After observation different platforms and tools, it was obvious there is no speech recognition from Uzbek to English and vice versa. Google provide this recognition for other languages such as Russia, German, however, there is no any microphone in order for speech recognition for Uzbek language. Similarly, Yandex translator also does not provide and detection. Hence, it is essential to do tools by using AI for Uzbek language. More precise, a tool that incorporates Uzbek cultural references, idiomatic expressions, and regional variations into the translation algorithms and speech recognition. This involves databases or modules specifically designed to comprehend cultural nuances, ensuring more culturally sensitive and accurate translations. It is developed for tailoring to Uzbek that prioritizes understanding context and nuances within sentences.

Conclusion

Striving for the advancement of AI translation tools designed for Uzbek signifies a significant leap toward facilitating seamless cross-cultural communication. Prioritizing the development of linguistic resources, contextual understanding, and cultural sensitivity within these tools is pivotal. By fostering collaborative platforms, utilizing advanced neural methodologies customized for Uzbek, and ensuring grammatical precision, these enhancements aim to elevate Uzbek translations to new levels of accuracy and cultural resonance. This concerted effort towards refining AI translation for Uzbek brings us closer to a future where language barriers diminish, fostering deeper connections and mutual understanding among Uzbek speakers and the global community.

References:

1. Abror, Sh. (2023). Recent advancements in Uzbek NPL tools. *Kompyuter lingvistikasining zamonaviy texnologiyalari- CTCL"2023".* Ilmiy amaliy anjuman. www.myscience.uz.

2. Bughin, Jaques, Manyika James. 2018. "The promise and challenge of the age of artificial intelligence". McKinsey Global Institute, May. https://www.mckinsey.it/idee/the-promise-and-challenge-of-the-age-ofartificial-intelligence 30/10. (Last accessed 13/10/2018).

3. Forcada, Mikel L., 2017. "Making sense of neural machine translation", in Translation Spaces (6)2 pp. 291–309 https://www.dlsi.ua.es//~mlf/docum/forcada17j2.pdf (Last accessed 29/10/2018).

4. Galati, Gabriele & Riediger Hellmut, 2017. "Traduttori Umani o Traduzione Automatica Neurale?". In Laboratorio Weaver, 24/06. http://www.fondazionemilano.eu/blogpress/weaver/2017/06/24/traduzioneautomatica-neurale/ (Last accessed 09/11/2018).

5. Kai, S. & Yue, Z. & Weihua, L (2018). Improved English to Russian translation by Neural suffix Prediction. https://dioi.org/1048550/arXiv 1801.03615

6. Munday, Jeremy. (2012). Introducing Translation Studies: Theories and Applications. ResearchGate.
https://www.researchgate.net/publication/317451115_Jeremy_Munday_Introducing Translation_Studies_Theories_and_Applications.

7. Rudy, S. & Christanta, R. (2020). Problems in the application of google translate as a learning media in translation.384-389. https://dioi.org/10.30743/ii.v4i2.2893

8. Snell-Hornby, Mary. 2006. The Turns of Translation Studies - New paradigms or shifting viewpoints, Amsterdam; Philadelphia: John Benjamins Publishing Company.