

«Сейфуллин оқулары – 18: « Жастар және ғылым – болашаққа көзқарас» халықаралық ғылыми -практикалық конференция материалдары = Материалы международной научно-практической конференции «Сейфуллинские чтения – 18: « Молодежь и наука – взгляд в будущее» - 2022.- Т.І, Ч.IV. - Б. 14-15

МАШИНАЛЫҚ АУДАРМА ЖАДЫНЫҢ ТЕХНОЛОГИЯСЫ

*Габдуллина Ж.А., 2 курс магистранты
С.Сейфуллин атындағы Қазақ агротехникалық университеті, Нұр-Сұлтан қ.*

Кіріспе.

«Машиналық аударма» термині көп мағыналы. Мәтіндермен онлайн жұмыс істеуге және жылдам аударуды жүзеге асыруға мүмкіндік беретін машиналық аударма (МА) құралдары ғаламтор пайдаланушылары үшін байланыс құралы ретінде қызмет етеді. Осындай құралдардың бірқатар кемшіліктері бар – қазіргі уақытта қолданыстағы платформалардың ешқайсысы жылдамдық, аударманың дұрыстығы тұрғысынан оңтайлы емес [1]. Аударма – адам қызметінің күрделі де көп қырлы түрі. Аударма процесі де сөйлеу шығармасындағы ақпаратты басқа тілдің көмегімен беру процесі болып табылады. Аударма барысында тілдер ғана емес, мәдениеттер де соқтығысады [2]. Мақалада аударма жадының технологиясы, оның алгоритмі, машиналық аудармаға кедергі тудыратын факторлар қарастырылады.

Машиналық аударма өзінің ұзақ пайдалану тарихында көптеген интерпретацияларға ие болды. Аударма жадының технологиясы (Translation Memory) аударма ережелерін қолданады және кіріс құжатын тұрақты жинақталған аударма базасындағы мәтіндермен салыстырады [3]. Сәйкестіктерді табу арқылы бағдарлама бұрын мақұлданған нұсқаны ұсынады.

Жүйедегі кіріс сөйлем бойынша іс-әрекеттердің стандартты алгоритмі келесідей[4]:

- морфологиялық талдау – сөйлем мүшелерін іздеу, кіріс сөз формаларын анықтау;
- берілген пән саласына қатысты идиомаларды, фразеологиялық бірліктерді іздеу және оларды әрі қарай талдаудан шығару;
- синтаксистік талдау – құрылымын талдау, сөйлем мүшелері – субъект, предикат, объект, жағдаяттарды табу – лексикалық талдау – бір мағыналы кірме сөздерді (лексемаларды) көп мағыналы сөздерден (бірнеше аударма баламасы бар) ажырату;
- грамматикалық талдау – оқытылатын тіл деректерін ескере отырып, грамматикалық ақпаратты анықтауға;
- шығыс сөйлемнің синтезі (аударма).

МА жүйелері салалас құрмалас және сабақтас құрмалас сөйлемдердің көп саны бар мәтіндермен жұмыс істеуге жарамсыз. Бұл бағдарламалар қарапайым жалпы сөйлемдерді қолданумен сипатталатын және құрамында күрделі сөйлемдер жоқ техникалық құжаттама, тұтынушы нұсқаулары, ресми сипаттамалар және т.б. сияқты формальды мәтіндерді аудару үшін сәтті пайдаланылуы мүмкін [5].

Машиналық аударманы қолданудың мақсаты жоғары сапалы аударманы алу немесе бастапқы мәтіннің мағынасын жеткізу. Машиналық аударма тілді локализациялау саласында белсенді қолданылады, мұнда аударуды қажет ететін ақпарат көлемі адам аудармашыларының нақты мүмкіндіктерінен асып түседі. Тілді локализациялау процесі тек аударма ғана емес, сапалы аударма сапалы локализацияның кілті болып табылады [6].

Машиналық аудармаға кедергі келтіретін негізгі факторларға мыналар жатады [7]:

- лексикалық және грамматикалық болуы мүмкін тілдік түсініксіздік;
- бастапқы тілде және аударма тілінде айтарлықтай ерекшеленуі мүмкін күрделі синтаксистік құрылымдардың болуы;
- сөйлемдегі сөздердің орын тәртібіндегі айырмашылықтар (тікелей/кері, қатан/еркін);
- мәтінде анафоралық сілтемелердің болуы;
- мағынасын сөзбе-сөз аудару арқылы жеткізуге болмайтын идиомалардың болуы;

Пайдаланылған әдебиеттер тізімі:

1. Борисова И. А. Коммуникация между интернет-пользователями—носителями различных языков // Вестник МГЛУ. 2013. №13 (673) С.28-34
2. Ааматов А. М. К вопросу машинного перевода: энтропия языковой системы и способы ее преодоления // Вестник ЛГУ им. А.С. Пушкина.2008. №2 (13) С.71-90.
3. Ахманова О. С. Словарь лингвистических терминов. М., 1969.
4. Карасев И. В., Артюшина Е. А. Системы машинного перевода // Успехи современного естествознания. 2011, №7, С.117-118.
5. Молчанов А. Статистические и гибридные методы перевода в технологиях компании ПРОМТ. М., 2013.
6. Zhechev V., Genabith V. Seeding Statistical Machine Translation with Translation Memory Output through Tree-Based Structural Alignment. //Proceedings of the 4th Workshop on Syntax and Structure in Statistical Translation, Beijing, China, 2010, P. 43—51.
7. Шевчук, В. Н. Информационные технологии в переводе. Электронные ресурсы переводчика. М., 2013.