

РАСПОЗНАВАНИЕ ЯКУТСКОГО ЯЗЫКА

*Захаров Тимур Захарович,
Подорожная Екатерина Сергеевна*

Студент (магистр), Студент (магистр)

*Институт математики и информатики СВФУ имени М. К. Аммосова,
Якутск, Россия*

E-mail: timu.zaxarov@gmail.com

Научный руководитель — Степанов Сергей Павлович

Автоматическое распознавание речи для малоресурсных языков представляет значительную сложность в сфере обработки естественного языка. В данном исследовании оценивается эффективность применения модели Whisper для распознавания якутской речи с особым фокусом на роль энкодера в повышении точности распознавания.

Якутский язык, как представитель тюркской языковой группы с ограниченными вычислительными ресурсами, требует специализированных подходов к созданию систем распознавания речи. В работе использовалась энкодер-декодерная архитектура Whisper Large-V3, адаптированная для якутского языка. Эксперименты проводились в трех конфигурациях: с полной моделью, с замороженным энкодером и с добавлением шумов в обучающие данные.

Результаты исследования продемонстрировали критическую важность энкодера для качественного распознавания якутской речи. При использовании замороженного энкодера показатель Word Error Rate (WER) составил 42%, тогда как без энкодера этот показатель увеличился до 45%. Функция потерь с энкодером опустилась ниже 0.1, а без него — увеличилась до 0.2. Эксперименты также выявили существенное влияние качества данных на эффективность модели: низкое качество данных на третьей эпохе обучения привело к заметному ухудшению результатов.

Полученные результаты подтверждают, что энкодер выполняет ключевые функции в модели Whisper при распознавании якутской речи: извлечение акустических признаков, моделирование контекста и обеспечение устойчивости к шуму. Исследование вносит вклад в развитие языковых технологий для малоресурсных языков и открывает перспективы для дальнейшей оптимизации гиперпараметров и расширения обучающей выборки.