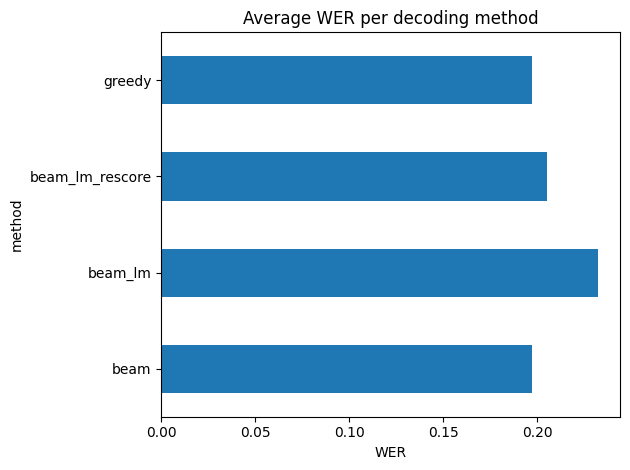
В ходе эксперимента была проведена оценка качества распознавания речи с использованием различных стратегий декодирования модели Wav2Vec2. В качестве языковой модели использовались как маленькая, так и большая N-грамм модели, доступные на сайте OpenSLR.

Цель эксперимента — определить, как использование языковой модели, изменение параметров beam search (beam\_width, alpha, beta), а также применение повторной переоценки (rescore) влияют на точность распознавания.

# Средняя WER по методам декодирования



# Выводы

- Метод beam search с повторной переоценкой (beam\_lm\_rescore) дал один из наименьших показателей WER, однако его преимущество по сравнению с обычным beam search и greedy decode не столь значимо, как ожидалось.

- Использование языковой модели в методе beam\_lm показало нестабильные результаты и, в некоторых случаях, даже ухудшение качества.

- Простая стратегия greedy показала сопоставимую точность, но при этом была значительно быстрее других методов.

- Увеличение beam\_width и параметров alpha/beta улучшает качество, но требует больше времени на обработку.

# Заключение

В большинстве случаев, разумный компромисс между точностью и производительностью достигается при использовании beam search с небольшим beam\_width и без языковой модели. Однако в критичных задачах распознавания речи, где важна точность, целесообразно использовать метод beam\_lm\_rescore с подобранными параметрами и мощной LM моделью.