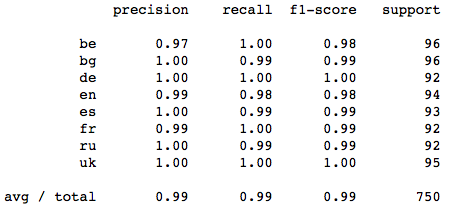
**Отчёт по определению языка в тексте**

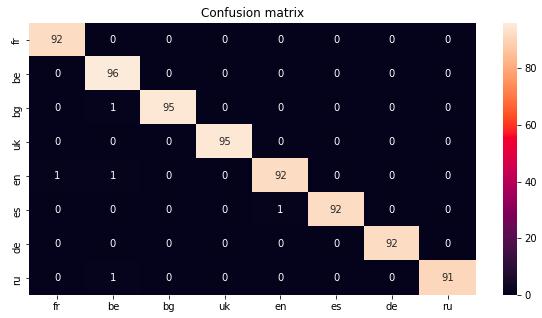
Корпусный метод работает лучше. У него процент ошибок меньше и, соответвенно, точность выше, чем у n-граммного. На confusion matrix видим, что путаются языки с одинаковым алфавитом (кириллическим / латиницей). Корпусный (словарный) метод лучше: мы отфильтровали повторяющиеся n-граммы, то есть в список попали менее частноные n-граммы (стоит учесть, что в языках больше совпадений n-грамм, чем слов). Возможно, это сказалось на точность второго метода.

**Процент ошибки у корпусного метода: 0.006666666666666667**

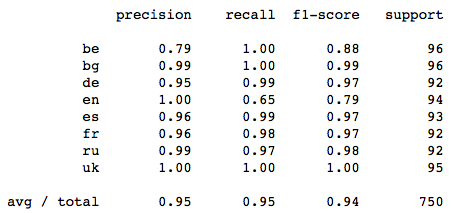
**Процент ошибки у n-граммного метода: 0.05333333333333334**

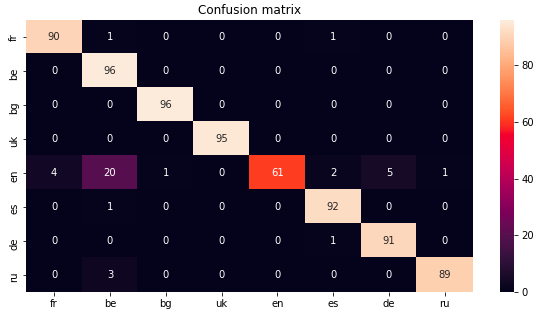
**Корпусный (словарный) метод**

****

****

**n-граммный метод**

****

****