raw corpus

prepare

abkhazia corpus

split

test set   **train set**

acoustic model   language model

forced alignment   decoding