Pareto front of BLEU score vs runtime

