

1. Зачем нужен dropout? Как его использование в модели с архитектурой CNN сказывается на качестве классификации и почему?
2. Результаты ConvNet -- на каких датасетах качество предсказания выше, чем у других моделей, на каких -- ниже? Есть ли какие-нибудь закономерности, и если да, то какие?
3. Чем LSTM отличается от Simple RNN?
4. Вариации LSTM, их отличия от стандартной LSTM
5. Чем отличаются LSTM и Bi-LSTM? Для решения какой проблемы была придумана Bi-LSTM?
6. В чем основная идея C-LSTM?