

1. Как раскладывается ошибка (excess error) с учетом ошибки оптимизации?
2. Как качественно зависит время обучения от точности оптимизации ρ , сложности семейства \mathcal{F} , количества данных n ?
3. В чем заключаются различия задачи обучения на малых и больших объемах выборки?
4. Какие характеристики функционала ошибки позволяли оценить скорость сходимости методов второго порядка в докладе?
5. Из каких соображений авторы работы предлагают выбирать точность оптимизации?