

НИС машинное обучение и приложения, 2016/2017

1. Назовите пару подходов к решению задачи определения схожести документов кроме Word Mover's Distance. В чем они заключаются?
2. Что такое обучение метрики? Зачем оно используется?
3. Какой функционал минимизирует NCA (в общем случае)? Как он связан с KL-дивергенцией?
4. Как выглядит оптимизационная задача для Supervised Word Mover's Distance?
5. Зачем мы используем энтропийный регуляризатор?