

## Отзыв

на диссертационную работу Сафоновой А.В.  
«Анализ пользовательского поведения для уточнения релевантности документа  
в Web - поиске»

Диссертационная работа Сафоновой А.В. посвящена разработке методов ранжирования документов при Web – поиске по факторам, определяемым исходя из данных о поведении пользователя на странице ответа поисковой системы. Учет поведения пользователей позволит провести реальную оценку релевантности результатов поисковых запросов и внести в поисковые алгоритмы дополнительные ранжирующие функции, что позволит повысить качество поисковых систем. Поэтому разработка методов ранжирования документов на основе пользовательского поведения является актуальной задачей.

Для решения поставленной задачи автор провел анализ исходных данных, содержащихся в пользовательских поисковых логах, и выявил ранжирующие факторы, сконструированные на их базе. При этом показано, что к наряду с известными факторами следует учитывать также факторы, связанные со степенью смещенности запроса, факторы региональности и skip – факторы. В целом предложено использовать для построения ранжирующей функции 109 факторов. Особенность задачи построения ранжирующей функции в данном исследовании позволяет ее трактовать как задачу построения классификатора при очень малой обучающей выборке. В работе классификатор строится с помощью метода опорных векторов, а сама ранжирующая функция при этом ищется как решение оптимизационной задачи выбора функции из семейства ранжирующих функций, полученных построенных на случайных подмножествах обучающей выборки. Точность полученной ранжирующей модели оценивалась с помощью 3-х различных мер – AUC, nDGG и MAP, применение которых позволило сравнить результаты ранжирования, полученных от различных алгоритмов обучения.

Реализация предложенного метода ранжирования проводилась при использовании библиотек SVM-Light и SVM-Rank. С помощью разработанного программного обеспечения было проведено большое количество экспериментов на реальном пользовательском логе объемом 340 миллионов записей и обучающей выборке объемом 33000 пар запрос – документ с экспертной оценкой релевантности. Результаты экспериментов показали конкурентоспособность подобного подхода по сравнению с другими методами оценки релевантности по пользовательскому поведению.

Считаю, что диссертационная работа Сафоновой А.В. заслуживает оценки “Отлично”, результаты могут быть рекомендованы к публикации, а ее автор может быть рекомендован для поступления в аспирантуру.

Научный руководитель  
доцент факультета ПМ-ПУ  
к.т.н.



В.М.Гришкин