

УТВЕРЖДАЮ

Первый проректор – проректор по научной
работе Московского физико-технического
института (государственного университета),
доктор технических наук, профессор



Горшков
Олег Анатольевич

«05» декабря 2014 г.

ОТЗЫВ

ведущей организации на диссертацию **Шаповалова Романа Викторовича**
«Методы структурного обучения в задачах совместной разметки»,
представленную на соискание ученой степени кандидата физико-математических наук
по специальности 01.01.09 — дискретная математика и математическая кибернетика

Актуальность темы

Тема диссертации относится к области распознавания образов. Несмотря на то, что задачи в этой области ставятся с середины прошлого века, научное сообщество ещё не научилось решать многие из них. В последние несколько лет, однако, прогресс в их решении заметно ускорился. С одной стороны, стали доступны обучающие выборки большого размера: например, в задачах компьютерного зрения выборки могут состоять из тысяч, или даже миллионов, фотографий. С другой стороны, вычислительные мощности, доступные исследователям, выросли. При этом классические методы распознавания образов не всегда позволяют использовать эти преимущества. Таким образом, необходима разработка нового математического аппарата для тонкой настройки статистических моделей по выборкам очень большого размера.

Подобные модели позволяют решать более сложные задачи, чем одномерная классификация. В диссертации дана характеристика класса задач так называемой совместной разметки, при решении которых можно использовать зависимость между предсказываемыми случайными величинами. Описаны методы, которые способны эффективно решать подобные задачи за счёт эксплуатации обучающих выборок большого размера и современных вычислительных мощностей. Отдельно стоит отметить проблему аннотации обучающей выборки такого размера: она требует значительных человеческих ресурсов. В работе рассматриваются методы, позволяющие настраивать параметры модели по неполной аннотации, получение которой значительно проще.

Научная задача и подход к решению

Диссертация посвящена решению задачи совместной разметки. Формально, в этой задаче по признаковому описанию объекта необходимо предсказывать его разметку — многомерный вектор категориальных переменных. Для этого по обучающей выборке, состоящей из пар признаков и разметок, строится некоторая статистическая модель, при этом предполагается, что условное распределение на разметки при условии признаков не изменяется со временем, таким образом, модель можно использовать для разметки вновь поступающих данных. Примером такой задачи служит разметка фотографических изображений с целью определения категорий объектов и их одновременной сегментации.

В работе исследованы два подхода к решению поставленной задачи. Первый основан на модели структурного метода опорных векторов, являющейся обобщением алгоритма SVM на случай многомерных целевых переменных. В нём каждому из вариантов разметки ставится в соответствие значение функционала энергии. На этапе предсказания находится разметка, минимизирующая энергию. Функционал линейно параметризуется, и на этапе обучения находятся значения параметров, такие что для объектов обучающей выборки минимум энергии достигается на разметках, близких к верным.

Второй подход основан на применении идеи последовательной классификации для учёта контекстуальных зависимостей. На этапе предсказания выполняется несколько итераций независимой классификации компонент разметки, при этом, начиная со второй итерации, в признаковое описание включается результат предыдущей итерации. Эта схема может быть реализована вычислительно эффективно.

При обучении структурного метода опорных векторов решается задача квадратичной оптимизации. На итерациях алгоритма вызываются дискретные подзадачи минимизации энергии. Для этого в работе используются недавние результаты из области дискретной оптимизации (Kohli et al., 2007; Lempitsky et al., 2009; Delong et al., 2012). При обучении последовательной классификации используется алгоритм индукции рандомизированных деревьев (Breiman, 2001).

Основные результаты и их новизна

В рамках диссертационного исследования Р.В. Шаповаловым были получены следующие результаты.

1. Разработан метод обучения нелинейного структурного метода опорных векторов. Метод успешно применён для настройки параметров неассоциативной Марковской сети. Приведено обобщение этого метода на случай обучения с латентными переменными и показана теоретическая связь полученного алгоритма с ЕМ-алгоритмом, используемым при вероятностном обучении Марковских сетей.

2. Предложена новая методика определения функций потерь при обучении структурного SVM по выборкам с различными типами слабой аннотации. Она опробована на различных типах слабой аннотации в двух прикладных задачах. Для каждого из типов аннотации выведены эмпирические функции потерь, являющиеся оценками стандартной эмпирической функции потерь. Наиболее близким является метод обучения структурного SVM по разнообразным данным (Kumar et al., 2011), который решает задачу в несколько стадий, аналогично жадному алгоритму. В отличие от него, в предложенном методе минимизируется одна целевая функция, а также используются другие типы аннотаций.
3. Разработан новый метод семантической сегментации на основе последовательной классификации, позволяющий учитывать контекстуальные зависимости в данных. В отличие от аналогичных подходов, таких как (Ross et al., 2011), предложенный метод использует априорную информацию о структуре пространственных зависимостей между элементами сцены.
4. Проведено экспериментальное сравнение предложенных методов с существующими на различных наборах реальных данных для задач компьютерного зрения и вычислительной лингвистики, показавшее превосходство предложенных методов по точности и скорости работы.

Достоверность полученных результатов

Достоверность теоретических результатов обусловлена подробным изложением доказательств, допускающим их проверку. Корректность экспериментов обусловлена подробным описанием протоколов, позволяющим воспроизводить эксперименты, и использованием стандартных наборов данных, допускающим сравнение с аналогами.

Значимость полученных результатов для практики

Практическая значимость работы Р.В. Шаповалова заключается в том, что разработанные методы улучшают применимость подходов к задаче совместной разметки, особенно в случае больших размеров обучающей выборки. Более конкретно, обучение по слабо-аннотированным данным позволяет использовать большие обучающие выборки, для которых получение полной разметки вручную было бы невозможно. Использование нелинейных моделей и неассоциативных марковских сетей позволяет настроить более гибкую модель, то есть, более точно описывающую генеральную совокупность при наличии достаточного количества обучающих данных. Методы на основе последовательной классификации позволяют повысить скорость работы методов как на этапе обучения, так и при предсказании.

Полученные результаты могут использоваться в научных организациях и компаниях, занимающихся исследованиями в области распознавания образов, а также прикладными разработками в области компьютерного зрения, вычислительной лингвистики, биоинформатики. Методы могут быть особенно интересны компаниям, занимающимся анализом больших объёмов веб-данных, таким как Яндекс, Google, Amazon, Facebook, Baidu.

Замечания

1. В обзоре литературы отсутствуют упоминания основополагающих трудов В.Н.Вапника, М.И.Шлезингера по теории распознавания образов.
2. При проведении экспериментов по категоризации текстовых документов не проводится сравнение с существующими методами. За основу алгоритма был взят метод, представляющий набор меток с помощью полносвязной Марковской сети, при этом этот выбор не был обоснован.
3. В разделе 2.2.4 недостаточно подробно описаны алгоритмы дискретной оптимизации, используемые при обучении структурного SVM с аннотациями типа «зёрна».

Приведённые замечания не влияют на общую положительную оценку диссертации.

Заключительная оценка

Диссертация Р.В. Шаповалова является законченной научно-исследовательской работой, имеющей высокую теоретическую и практическую ценность, а её тема является актуальной. Полученные результаты являются новыми и достоверными. Результаты развивают математическую теорию распознавания образов, а методология включает в себя приёмы из областей математической статистики, дискретной и непрерывной оптимизации, что обосновывает соответствие специальности 01.01.09 — дискретная математика и математическая кибернетика.

Полученные результаты опубликованы в научной печати и докладывались на российских и международных конференциях. Автореферат полностью и точно отражает содержание диссертации. Диссертационная работа «Методы структурного обучения в задачах совместной разметки» полностью соответствует требованиям Положения о присуждении учёных степеней, предъявляемым к диссертациям на соискание степени кандидата физико-математических наук, а соискатель заслуживает присуждения учёной степени кандидата физико-математических наук по специальности 01.01.09 — дискретная математика и математическая кибернетика.

Отзыв на диссертацию составлен доктором физико-математических наук, профессором Буниной Еленой Игоревной и обсуждён на заседании кафедры анализа дан-

ных факультета инноваций и высоких технологий Московского физико-технического института (государственного университета) «15» ноября 2014 г., протокол № 4.

Заместитель заведующего кафедрой анализа данных
факультета инноваций и высоких технологий МФТИ,
доктор физ. – матем. наук, профессор



Бунина
Елена Игоревна

Почтовый адрес: 141700, Московская область, г. Долгопрудный, Институтский пер., 9

Телефон: 8 (495) 408 78 81

Адрес электронной почты: helenbunina@yandex.ru

Организация – место работы: федеральное государственное автономное образовательное учреждение высшего профессионального образования «Московский физико-технический институт (государственный университет)»