

Десять лекций по статистическому и структурному распознаванию

МИХАИЛ ИВАНОВИЧ ШЛЕЗИНГЕР
Национальная Академия наук, Украина, Киев

ВАЦЛАВ ГЛАВАЧ
Чешский Технический Университет, Чешская республика, Прага

Введение

Предисловие к русскому изданию

Известно, что введение к книге пишется после того, как книга уже написана, и то, что получилось, можно представить как задуманное с самого начала. В этом отношении наша книга находится в лучшем положении. Мы пишем это введение после того, как книга уже вышла сначала в чешской, а затем в английской версии. Таким образом, мы уже имеем некоторые представления о том, что получилось, а что еще нет.

Во введении к научной работе автор как бы извиняется за ее написание, доказывает, что она кому-то нужна, и так или иначе избегает признания, что выполненная работа нужна прежде всего ему самому. Мы не видим в таком признании ничего зазорного. Наоборот, если какая-то работа делается для себя, то она окажется нужной еще кому-то скорее, чем если бы она делалась только для продажи. Ведь не зря говорят: сделайте, как для себя, когда надеются на добросовестную работу. Мы написали эту книгу прежде всего потому, что она была крайне необходима нам самим. При этом неизбежно достигаются еще и какие-то другие цели. Укажем некоторые из них.

1. Мы оба преподаем различные разделы распознавания образов в Национальном техническом университете Украины «Киевский политехнический институт», Чешском техническом университете в Праге, в Международном научно-учебном центре информационных технологий и систем НАН Украины и МОН Украины в Киеве. Отдельные курсы нами были прочитаны в Университете имени Масарика в Брно (Чехия) и Дрезденском техническом университете. Содержание этой книги в значительной степени пересекается с содержанием наших лекций.

Как и в любой студенческой группе, в наших группах оказывалась часть студентов, которые отсутствовали на лекциях по причинам, которые им казались уважительными. В конце семестра они вынуждены были освоить учебный материал, располагая только данной книгой. Многие из них справились с этим вполне успешно. Написанная книга, по-видимому, вполне пригодна для самостоятельного изучения основ распознавания образов, начиная с самых азов.

2. Мы оба работаем в коллективах, каждый в своем, в которых профессионально исследуются проблемы распознавания образов. Время от времени к нам приходит студент старшего курса или недавний выпускник и выражает желание работать в нашей группе. Распознавание образов не входило в вузовский курс его обучения, но он много слышал об этой области знаний, и она показалась ему привлекательной. Он уверен, что сможет плодотворно работать в этой области. Однако мы в этом не столь категорично уверены, потому что знаем, насколько сильно отличается проза науки о распознавании от поэтических представлений об искусственных устройствах, наблюдающих и понимающих окружающий мир так же, как человек или другие живые существа. Различие здесь приблизительно такое же, как между любовью к музыке и желанием овладеть фортепианным мастерством. Мы обязаны как можно раньше показать молодому исследователю «кухню» распознавания и определить, действительно ли его привлекает наука о распознавании настолько, чтобы затратить на нее те усилия, которые она требует. Или эта наука — нечто другое, отличное от того, что он ожидал. Такое ознакомление не должно предполагать предварительных знаний о распознавании, но должно проводиться достаточно энергично, чтобы быстро вывести будущего исследователя на современный уровень проблем. Тем не менее этот процесс требует определенных затрат, которые в ряде случаев, увы, оказываются напрасными: мы вместе с нашим молодым коллегой приходим к выводу, что существует очень много других областей информатики, которые для него более привлекательны, чем распознавание образов. Сейчас, когда у нас есть эта книга, потерь такого рода стало значительно меньше.

3. Книга написана не только для наших молодых коллег, но и для исследователей, имеющих опыт в распознавании образов. Если бы мы адресовали книгу нашим опытным коллегам, то наиболее подходящим названием для нее было бы: «Неизвестное об известном в распознавании образов». Мы оба работаем в распознавании образов довольно давно. Естественно, мы получили некоторые новые результаты, что само по себе еще не является поводом для написания книги. Существенно, что некоторые из результатов удалось получить только после критического анализа ряда общепринятых представлений и понимания того, что эти представления неоправданно сужают проблему, а иногда просто ошибочны. К счастью, все эти предрассудки, несмотря на свою распространенность, не укоренились слишком глубоко, так как наука о распознавании еще сравнительно молода и не приобрела жесткие рамки. Преодоление этих предрассудков не вызовет какой-то болезненной ломки, и рано или поздно они отомрут сами собой. Мы написали эту книгу, чтобы это произошло как можно раньше.

4. И наконец, мы укажем ту цель, которую ставили до написания книги, но о которой не можем сейчас с уверенностью сказать, что она этой книгой достигается.

Мы считаем, что к настоящему времени уже достаточно четко сформировался круг прикладных проблем распознавания образов. Область приложения распознавания образов разумно сужена и сейчас в нее уже не вклю-

чают, например, машинную игру в шахматы или решение проблем управления, как это было при ее зарождении полвека назад. Значительно менее четко сформирована чистая научная проблематика на уровне абстрактных формальных понятий, не связанных с тем или иным конкретным приложением. Мы считаем, что наука о распознавании находится сейчас на таком этапе, что для ее прогресса требуется анализ абстрактных математических моделей в большей степени, чем дальнейшее наращивание списка возможных приложений. Это требует привлечения в нашу науку новых сил, с несколько иным, если можно так сказать, научным темпераментом. К сожалению, этому не способствуют установившиеся традиции внутри нашего же распознавания. Считается, например, неприличным опубликовать статью по проблеме распознавания, в которой во введении не указано конкретное приложение, а сама статья не завершается какой-то картинкой. Этим достигается лишь то, что плодотворная научная идея оказывается безнадежно замаскированной конкретным приложением и становится мало доступной тем профессионалам, которые работают вне распознавания, но могли бы обогатить эту идею. Мы все как будто прикладываем специальные усилия к тому, чтобы нас как можно меньше поняли наши коллеги, работающие за пределами распознавания. Более того, мы будто стараемся, чтобы специалист по распознаванию, скажем рукописных текстов, ни в коем случае не понял работу по стереозрению.

Область возможных приложений распознавания образов исключительно обширна и многообразна. Однако мы сможем говорить о распознавании образов, как о науке, только тогда, когда научимся решать прикладные задачи оптом, а не в розницу. Для этого необходимо несколько отдалиться от прикладных задач и взять на себя смелость сформулировать и исследовать задачи распознавания в чистом виде, не загромождая их сущность теми или иными подробностями прикладного характера.

Этой книгой мы хотим выразить наше глубокое уважение и благодарность нашим учителям — профессору В.А.Ковалевскому и профессору З.Котеку. Если появление нашей книги — это заслуга, то это прежде всего их заслуга. Мы будем считать величайшей честью для себя, если наши учителя в число своих достижений включат и эту книгу.

Работа над всеми версиями книги — чешской, английской и русской — продолжалась значительно дольше, чем мы ожидали. Все это время нам оказывала неоценимую помощь Т.Н.Барабанюк и коллега В.Зика.

Авторы благодарны друг другу за сотрудничество, которое вопреки географической удаленности, различию в возрасте, научном темпераменте и языке удалось довести до сегодняшнего результата.

М.И.Шлезингер, В.Главач, сентябрь 2004

Предисловие к английскому изданию

Современная наука о распознавании характеризуется таким разнообразием методов и подходов, что создается впечатление о бесчисленных, не связанных друг с другом островках, рассеянных в океане нашего незнания. В

этом океане трудно ориентироваться и легко заблудиться. В данной монографии показано, что некоторые группы известных в настоящее время задач допускают общую формулировку. Таким образом обнаруживается родство задач, которые до сих пор казались различными и решались независимо друг от друга. Эти частные задачи уже перестают казаться разрозненными островками, а обретают форму по крайней мере архипелага с ясно видимыми путями перехода от одного островка к другому.

Построение общей схемы дает, однако, значительно больше, чем просто ориентацию в известных методах. Представление известных методов, как частных случаев единой общей схемы, позволяет обнаружить и другие, ранее не замеченные, но важные частные случаи.

Два обширных направления в современном распознавании — статистическое и структурное — изложены в книге с единой методологической точки зрения. Исследование задач структурного распознавания с точки зрения статистического позволило сформулировать и решить новые задачи структурного распознавания и, как следствие, понять ограниченную применимость известных методов. Помимо этого моста, объединяющего два крупных направления в современном распознавании, обнаружены взаимосвязи и отдельных задач внутри каждого из этих направлений. Это также привело к неожиданным результатам.

Монография адресована как нашим коллегам, которые уже имеют опыт в распознавании образов, так и будущим коллегам, которые нуждаются во введении в проблему. Материал книги излагается, начиная с азов, так что для ее чтения не требуется каких-либо предварительных знаний. Требуется лишь желание эти знания приобрести. Для чтения большей части книги желательны навыки чтения не очень сложных математических текстов, так как основные результаты изложены в виде математически точных утверждений и их формальных доказательств. Ключевые понятия и рекомендации объясняются и иллюстрируются с помощью доступных интеллектуальных упражнений.

Выбранный стиль изложения, можно надеяться, делает книгу доступной для студентов и аспирантов в самом начале их научной карьеры. Именно таким читателям мы старались показать на конкретных примерах все прелести, парадоксы и ловушки распознавания. Каждая лекция завершается ее обсуждением с одним из лучших наших студентов. Дискуссии с ним позволили углубить понятия, введенные в лекции, сделать их более прозрачными, уменьшить вычислительную сложность введенных в лекции процедур. Когда это было уместно, мы указывали нашему юному коллеге на ошибочные представления, которые встречаются в распознавании образов, увы, значительно чаще, чем хотелось бы.

Лекции построены в стиле, в котором принято излагать математические дисциплины. Это значит, что основные понятия формально определены, а результаты сформулированы в виде теорем, для которых приводится их полное доказательство. Задачи распознавания и идеи их решения, изложенные в таком виде, не заслоняются ненужными подробностями того или иного приложения. Мы надеемся, что такой стиль изложения привлечет в

нашу науку исследователей с математическим складом ума. Наука о распознавании нуждается в таких исследователях и заслуживает их внимания. Большая часть современных публикаций, к сожалению, не способствует их привлечению, так как вместо формулировки решаемой математической задачи читатель находит в ней формулировку задач криминалистики, или медицины, или генетики, или экологии, или других. Широкая область прикладной применимости распознавания образов достаточно хорошо известна и в силу этого уже привлекает внимание специалистов самых разнообразных областей практики. Вместе с тем научное содержание решаемых нами проблем значительно меньше известно нашим коллегам, работающим в других областях информатики вне распознавания образов.

Сказанное никак не противоречит тому, что данные лекции написаны практиками и для практиков. Для представленных в лекциях алгоритмов точно описана их область определения и поведение. Это позволяет использовать их нашими коллегами-практиками как строительные блоки при построении сложных прикладных программных комплексов. Ряд алгоритмов, описанных в лекциях, реализован в виде toolbox-ов, работающих в MATLAB. Программное обеспечение доступно по WWW-адресам, указанным в книге. Программы написаны нашими студентами из Чешского технического университета.

Оба автора работают в Восточной Европе (Украина и Чешская республика) и следуют научным традициям своих национальных школ. В книге содержатся ссылки на работы по распознаванию образов, выполненные в наших странах. Многие из них неизвестны в других странах мира. Таким образом, книга открывает перед широкой научной общественностью по крайней мере часть картины, представляющей распознавание образов в наших странах.

Мы заинтересованы в обратной связи от читателей. Без лишних сомнений пошлите нам сообщения по адресам (schles@image.kiev.ua, hlavac@fel.cvut.cz). Желаем Вам увлекательного и полезного чтения.

М.И.Шлезингер, В.Главач, май 2001

Письмо аспиранта Иржи Пехи до опубликования лекций

Глубокоуважаемый господин профессор Главач, Чески Крумлов, 25.11.1996

Я — аспирант первого года обучения на электротехническом факультете Чешского технического университета. Только сейчас я узнал, что в весеннем семестре 1996 года некий профессор с Украины прочитал на нашем факультете курс лекций по математической теории распознавания образов. К сожалению, я об этом узнал слишком поздно и поэтому не смог посетить эти лекции. Я сожалею об этом все больше и больше, потому что коллеги, посещавшие лекции, часто на них ссылаются. Мне сказали, что эти лекции не опубликованы, но мне пришла мысль, что, возможно, у Вас

остались какие-то записки лектора или его конспекты этих лекций. Если это так, то позвольте мне, пожалуйста, снять с них копию.

Я намерен выполнить свою аспирантскую работу в области распознавания образов и понимания изображений. Некоторый опыт в этом направлении я уже приобрел, и этот опыт, к сожалению, далеко не положительный. Я написал довольно сложную программу по распознаванию буквенных символов по их изображениям. Я обошелся без чтения литературы и без глубоких исследований. В основу работы я положил свои собственные соображения, которые представлялись мне естественными и разумными. Однако полученные результаты оказались настолько неожиданно плохими, что поначалу я подумал, что допустил массу ошибок при программировании своих идей. После тщательной отладки программ я убедился, что они написаны правильно. Следовательно, ошибочными оказались мои идеи, казавшиеся мне почти очевидными.

Теперь я вижу, что распознавание образов — это необозримое поле, в которое трудно войти, а войдя в него, легко заблудиться. Я попытался восполнить недостаток своих знаний в распознавании по учебникам и журнальным статьям. Некоторые публикации, которые я просмотрел, скорее популяризуют проблему, чем описывают ее. С более глубокими публикациями у меня серьезные проблемы. При их чтении у меня создается ощущение, будто я читаю увлекательный роман, но читаю его не с самого начала, как положено, а откуда-то с середины, и мне остается только строить догадки, кто именно скрывается под тем или иным именем и как эти действующие лица связаны друг с другом. Конечно, для чтения этих работ мне не хватает каких-то элементарных, базовых знаний, на которые в данной публикации нет ссылки, так как они предполагаются общеизвестными. Это естественно, так как эти публикации не рассчитаны на таких новичков, как я.

От своих преподавателей я знаю, что математические средства современного распознавания значительно обширнее, чем тридцать лет назад, когда распознавание образов базировалось почти полностью на математической статистике в многомерных линейных пространствах. Современное распознавание является областью приложения результатов теории графов, формальных языков и грамматик, марковских цепей и полей, математического программирования, вычислительной и дискретной геометрии. Более того, новые алгебраические конструкции рождаются и внутри самого распознавания образов. Со всеми этими областями я знаком хуже, чем это нужно для того, чтобы в них свободно ориентироваться.

Кстати, о ссылках, которые я нашел в попавшихся мне публикациях. Чаще всего это были ссылки на малодоступные публикации того же автора, либо ссылки на солидные математические монографии, иногда многотомные. В них невозможно найти краткое и простое определение того элементарного понятия, которое необходимо для прочтения статьи. Я определенно прихожу к выводу, что в современной обширной литературе по теории и практике распознавания образов отсутствует монография, или учебник, или публикация какого-то иного рода, которая бы служила введением в

проблему. В такой книге я бы хотел найти определение наиболее важных понятий, формулировку базовых задач и сведения о тех математических средствах, которые необходимы для моей дальнейшей самостоятельной работы или, по крайней мере, для чтения научных статей по распознаванию. Книга должна знакомить с предметом при минимальных предположениях об осведомленности читателя. Короче говоря, мне не хватает книги, написанной специально для меня.

Как мне сказал мой научный руководитель, искать такую книгу-учебник-монографию совершенно тщетно. Мне дали совет, что до некоторой степени отсутствие такой книги восполняется лекциями профессора из Киева. Возможно, Вы будете так любезны и пришлете мне записи этих лекций, чтобы я смог снять с них копию.

Заранее искренне благодарю Вас.
Иржи Пеха

Письмо авторов аспиранту Иржи Пехе

Дорогой коллега Пеха,

Киев, 5 января 1997

в значительной мере мы согласны с Твоей оценкой современного состояния распознавания образов. Распознавание образов зародилось в 50-е годы и в 90-е годы переживает новый всплеск своей популярности. Область возможных применений здесь так широка, что трудно сказать, где только не попытаются применить распознавание образов. Эта область простирается от микромира до Космоса.

Естественно, такая популярность не может не вдохновлять каждого, кто профессионально работает в этой области. В то же время она вызывает оправданную тревогу. На распознавание образов опять, как и полвека назад, возлагают надежды, которые вряд ли когда-нибудь сбудутся. Распознавание образов потихоньку опять начинает обретать репутацию палочки-выручалочки, с помощью которой одним махом решаются практические задачи без их детального и изнурительного изучения. Такая репутация вызывает здравый скептицизм трезвых практиков и раздражение здравомыслящих исследователей. Для более доверчивых эта репутация станет источником серьезных разочарований, когда опять станет ясно, что одно лишь использование слов «распознавание», «нейронные сети», «искусственное мышление» и тому подобных вовсе не гарантирует решение практических задач.

Сейчас наступило самое подходящее время, чтобы произвести своего рода инвентаризацию распознавания образов и отделить действительность от сладких грез о чудодейственных способностях распознающих систем. Распознавание образов уже выросло из младенческого возраста, когда требовало к себе бережного отношения. Сейчас оно уже не нуждается в снисхождении, и его авторитет может строиться не на мифах, а на действительности.

Ты прав, что лекции, которые Ты пропустил, в какой-то мере содержат именно то, что Тебе нужно. Однако вряд ли Тебе будут полезны за-

писки о лекциях, которые сейчас у нас есть. В них содержатся формулировки основных понятий и теорем, и это лишь незначительная часть того, что происходило в аудитории. В лекциях содержалась краткая мотивация для рассмотрения того или иного раздела распознавания, затем критический (зачастую довольно строгий) анализ некоторых известных подходов и предостережение от всевозможных ловушек. Имеются в виду так называемые разумные решения, вроде бы основанные на здравом смысле, но на самом деле ошибочные. Именно после таких комментариев, дополнений к основному материалу, мы осознавали, как мало мы знаем даже о тех разделах распознавания, о которых думали, что разбираемся в них достаточно неплохо.

Мы неоднократно задумывались над тем, следует ли опубликовать все эти интересные, но не совсем академические соображения, и в свое время решили, что опубликовать их невозможно или, по крайней мере, очень трудно. Однако теперь мы все чаще сомневаемся в правильности нашего решения и начинаем думать, что публикация этих лекций будет полезной не только для Тебя. Сейчас мы приводим эти лекции к виду, удобному для чтения. Однако это будет несколько иная книга, чем Ты себе представлял. Ты хотел бы иметь книгу типа справочника, в котором бы отдельные главы были посвящены отдельным математическим дисциплинам, составляющим математический аппарат распознавания. Поначалу и мы думали написать именно такой справочник, но теперь изменили свои замыслы. Именно под влиянием Твоего письма мы поняли, что Ты нуждаешься не в математическом справочнике, а в чем-то ином.

Хотя Ты довольно скромно оцениваешь свои знания относительно математических средств современного распознавания, Ты правильно понимаешь, что речь идет, как правило, о базовых, т.е. самых элементарных сведениях, заимствованных распознаванием извне. Твои проблемы не в том, что эти сведения так уж сложны, а в том, что их много, они чересчур разнообразны и почерпнуты из разных математических дисциплин. До того, как эти понятия были перенесены в распознавание, они не соприкасались друг с другом. Новичка приводит в замешательство также то, что в рамках распознавания заимствованные понятия приобретают несколько иную окраску, а зачастую и иное название, чем в материнских математических дисциплинах. Современное распознавание состоит не только в использовании известных математических методов для решения той или иной практической задачи. Используемые в распознавании известные методы приходят в необычное для них соприкосновение, что приводит к новым понятиям, новым задачам и, в конечном итоге, к результатам, которые входят уже в науку о распознавании, а не в исходные математические дисциплины.

Лекции, которые мы сейчас приводим в порядок, посвящены проблемам распознавания, а не задачам линейного программирования, математической статистики или теории графов. Требуемые математические средства вводятся не изолированно друг от друга, а во взаимосвязи, в которую они должны вступить для решения той или иной задачи распознавания. Известные математические средства становятся таким образом составны-

ми частями единого механизма, который можно назвать математическим аппаратом распознавания. Нам кажется, что именно такого взгляда на распознавание Тебе сейчас не хватает и в этом состоят Твои затруднения.

К сожалению, сейчас мы только начали приводить лекции к виду, пригодному для опубликования. Работа потребует приблизительно два года. Мы понимаем, что Тебе не хотелось бы простаивать все это время, и предлагаем Тебе сотрудничество. Мы можем посылать Тебе тексты лекций по мере их написания. Было бы замечательно, если бы Ты сразу после получения части материала критически прокомментировал его, сообщил нам о своих идеях или вопросах, если они возникнут. Давай считать такую обратную связь необходимым условием нашего сотрудничества. Очередной материал Ты получишь только после того, как мы получим Твой анализ предыдущего.

Сейчас мы высылаем Тебе первую лекцию и с нетерпением ждем, что же получится в результате нашего сотрудничества.

Шлезингер Михаил Иванович, Главач Вацлав

Лекция 1

Байесовские задачи распознавания

1.1. Вводные замечания

Байесовская теория принятия решений является одним из блоков, из которых строится статистическая теория распознавания. В этой лекции мы введем основные понятия, сформулируем байесовскую задачу принятия решений и укажем наиболее важные общие свойства этих задач. Понятия, введенные на этой лекции, будут неоднократно использоваться на протяжении всего курса. Считается, что эти понятия общеизвестны, и мы были бы рады, если бы это было действительно так. В действительности же нередко встречаются рекомендации, которые явно противоречат выводам байесовской теории, хотя на первый взгляд представляются вполне естественными. Это свидетельствует о поверхностном, частичном понимании байесовской теории, а это хуже, чем полное незнание. Несомненно прав был тот, кто сказал, что лучше иметь дело с человеком, который не читает книг вообще, чем с тем, кто прочитал одну книгу.

Конечно, имеется много причин для разнобоя в понимании, что такое байесовская теория. Слишком уж знаменито имя Байес. Формула, по которой подсчитываются апостериорные вероятности, носит имя Байеса. Этим же именем называют решение о значении случайной величины в пользу наиболее вероятного ее значения. В современной прикладной статистике появилось понятие «байесовские сети» как модели сложного случайного объекта. Поэтому вовсе нередка ситуация, когда собеседники долго спорят о байесовском методе распознавания и только в конце выясняют, что говорят о разных вещах.

Мы настоятельно просим читателя, который не знаком с этими понятиями, не переходить к чтению последующих глав пока он не ознакомится с этой лекцией и комментариями Иржи Пехи в конце лекции. С этой же просьбой мы обращаемся к читателю, который знает байесовскую теорию: убедимся на всякий случай, что мы имеем в виду одно и то же.

Лекция 2

Небайесовские задачи распознавания

2.1. Жесткая ограниченность байесовского подхода

В первой лекции мы неоднократно подчеркивали большую общность байесовского подхода. Эта общность выражается в том, что байесовские задачи и основные свойства байесовских стратегий формулируются для множеств наблюдений X , состояний K и решений D с самыми разнообразными математическими свойствами. Очень желательно владеть всем тем богатством, которым обладает байесовский подход, и не отождествлять байесовский подход в целом с тем или иным его частным случаем. Мы уже знаем, что класс байесовских задач — это нечто большее, чем минимизация вероятности ошибочного решения.

При всей общности байесовского подхода существуют задачи, которые не могут быть выражены в его рамках. Это так называемые небайесовские задачи статистической теории решений. Их нужно знать, чтобы для каждого конкретного приложения выбирать наиболее уместную формализацию, а не наоборот. Исходную прикладную задачу не следует деформировать и неестественно втискивать в рамки наперед заданного формализма, пусть даже такого общепризнанного, как байесовский подход.

Байесовский подход ограничивается сразу же после введения тех понятий, которые лежат в его основе. Во-первых, это *функция потерь* $W: K \times D \rightarrow \mathbb{R}$. Во-вторых, это *совместные вероятности* $p_{XK}(x, k)$, $x \in X$, $k \in K$, которыми вводятся два дополнительных понятия: *априорные вероятности* $p_K(k)$ ситуаций $k \in K$, которые должны быть распознаны, и *условные вероятности* $p_{X|K}(x | k)$, $x \in X$, $k \in K$, наблюдений x при условии ситуации k . Рассмотрим более тщательно, как введение этих понятий сужает круг прикладных задач, которые могут формализоваться в рамках байесовского подхода. Для начала, в подразделах 2.1.1–2.1.3, это рассмотрение будет неформальным.

2.1.1. Функция потерь

Как только произнесены слова «минимизация математического ожидания потерь», так сразу ограничивается характер множества значений, которые может принимать штраф. Это должно быть полностью упорядоченное множество, на котором, кроме отношений $<$ или \geq , должны быть определены сложение штрафов и умножение их на вещественное число. Для байесовского подхода существенно, что штрафы измеряются вещественными числами, и это ограничение не представляется обременительным при абстрактном рассмотрении формальных моделей. Однако существует ряд приложений, где такая неестественность бросается в глаза и сразу же видно, как отождествление штрафа с вещественными числами деформирует прикладное содержание задачи самым недопустимым образом. Такие прикладные задачи не должны решаться в рамках байесовского подхода. Они должны формулироваться без использования понятия штрафа.

Здесь можно вспомнить, правда, не слишком серьезно, известного героя русского фольклора: повернешь налево — коня потеряешь, повернешь направо — меч потеряешь, пойдешь прямо — потеряешь любимую. Знание байесовского метода здесь мало поможет нашему герою, так как никаким вразумительным способом ему не удастся определить, что больше: сумма p_1 (лошадей) и p_2 (мечей) или p_3 (любимых). Само понятие суммы p_1 (коней) и p_2 (мечей) бессмысленно, так же, как не имеет смысла сумма 30 метров и 80 секунд.

В этом случае мы сталкиваемся с той нередкой ситуацией, когда функция потерь полностью задана, но она принимает значения, которые *не могут измеряться в одной системе единиц*. Вспомним, теперь уже чуть более серьезно, такую обширную область приложений, как диагностика сложного оборудования, когда на основании наблюдений следует определить, находится ли оборудование еще в исправном состоянии или в состоянии, которое уже опасно для окружающих. Каждая ошибочная диагностика связана с определенными потерями, которые известны. Но потери, когда исправное оборудование воспринимается, как опасное, совсем иного рода, чем потери при упущенной опасной ситуации. В первом случае потери составляют стоимость профилактических мероприятий, которые можно было бы и не проводить. Во втором случае потери могут быть совсем иные и не могут отождествляться ни с каким количеством профилактик. Например, может разрушиться что-то, что природа создавала миллионы лет. Несравнимость потерь от ошибок различного рода еще более выразительна в задачах, где объектом диагностики является человеческое общество или отдельно взятый человек, будь то медицинская диагностика или криминалистическая.

2.1.2. Априорные вероятности состояний

Чтобы сформулировать задачу распознавания в байесовских рамках, необходимо задать априорные вероятности $p_K(k)$ каждого состояния $k \in K$. На

самой поверхности лежит трудность, состоящая в том, что эти вероятности неизвестны. Однако при более тесном соприкосновении с прикладными задачами обнаруживаются и другие трудности, принципиально иного характера. Оказывается, что они принципиально не могут быть известны. Укажем три типа задач по степени увеличения рассматриваемых трудностей.

1. Состояние объекта случайно и априорные вероятности $p_K(k)$, $k \in K$, состояний известны. В этом случае задача может ставиться и решаться как байесовская.
2. Состояние объекта случайно, но априорные вероятности $p_K(k)$, $k \in K$, состояний неизвестны, так как объект распознавания еще не был достаточно хорошо исследован. Разработчик в этом случае имеет две возможности: а) он может сформулировать задачу не в рамках байесовского подхода, а как-то иначе, так, чтобы априорные вероятности состояний не требовались в качестве исходных данных задачи; б) он может более тщательно изучить объект распознавания, чтобы восполнить недостающие знания, необходимые для байесовского решения.
3. Состояние объекта не является случайным. В силу этого априорные вероятности $p_K(k)$, $k \in K$, не существуют и их невозможно получить никаким дополнительным исследованием объекта распознавания. В этом случае разработчик должен формулировать и решать свою задачу, как небайесовскую. Приведем пример такой ситуации, когда состояние объекта не случайно.

Пример 2.1 *Задача, не входящая в байесовский класс.* Пусть x — это сигнал, отраженный от наблюдаемого самолета. На основании этого сигнала нужно принять решение: свой ($k = 1$) или чужой ($k = 2$). Условные вероятности $p_{X|K}(x|k)$ могут очень сложно зависеть от наблюдения x . Тем не менее можно предположить, что по крайней мере существуют (хотя нам и неизвестны) такие числа $p_{X|K}(x|k)$, которые правильно отражают зависимость наблюдения x от состояния k . Что касается априорных вероятностей $p_K(k)$, то они неизвестны и не могут быть известны в принципе. Не существует такого числа α , $0 \leq \alpha \leq 1$, о котором можно было бы сказать, что это вероятность чужого самолета. Вероятности $p_K(k)$ просто не существуют, так как при сколь угодно длительном наблюдении частота появления чужого самолета не стремится к пределу, о котором можно было бы сказать, что это и есть вероятность. Иными словами, появление чужого самолета не является случайным событием. ▲

О неслучайном событии невозможно сказать, какова его вероятность, подобно тому, как невозможно сказать, какова температура звука или вкуса света. Такая характеристика, как вероятность, для неслучайного события просто не определена. Прикладные задачи, в которых требуется принять решение о неслучайном, но меняющемся состоянии объекта, не формализуются в рамках байесовского подхода. Они требуют иных теоретических

конструкций, в которых понятие априорная вероятность не используется вообще.

Нас всегда удивляла легкость, с которой незнание априорных вероятностей отождествляется с предположением, что все эти вероятности равны друг другу. В примере 2.1 это бы означало, что если неизвестна вероятность чужого самолета, то следует считать, что чужой самолет так же вероятен, как и наш. Но это не согласуется со здравым смыслом, даже если принять, что появление чужого события есть случайное событие: хоть нам и неизвестна вероятность чужого самолета, но ясно, что она отличается от вероятности своего. Мы не имеем веских доводов в пользу принятия гипотезы о равномерном распределении вероятностей в случае, когда это распределение неизвестно. Не исключено, что здесь имеет место одно из псевдорешений, когда логическое обоснование заменяется ссылкой на знаменитую личность, в данном случае на К.Шеннона, и общеизвестный факт, что энтропия максимальна именно при равномерном распределении вероятностей, хотя к рассматриваемым вопросам этот факт не имеет никакого отношения.

2.1.3. Условные вероятности наблюдений

Рассмотрим следующую прикладную задачу. Пусть X — множество изображений, на которых представлены буквы. Имя буквы k есть скрытый параметр изображения, а K — множество имен букв. Предположим, что буквы пишутся тремя людьми. Обозначим эту тройку Z , и пусть переменная $z \in \{1, 2, 3\}$ определяет, кто написал ту или иную букву. В этой ситуации возникают две различные задачи: распознавание буквы независимо от того, кто ее написал, и распознавание, кто написал букву, независимо от того, что это за буква. Мы будем говорить о первой задаче. В данной модели мы расширили круг основных понятий по сравнению с теми, которые были введены до сих пор. Мы сейчас уже говорим не о двух параметрах объекта: наблюдаемом признаке x и скрытом состоянии k , а еще и о третьем параметре z . Этот параметр скрыт от непосредственного наблюдения. Но в отличие от скрытого состояния k нас не интересует его значение. Целью распознавания является принятие решения о состоянии k , а не о значении параметра z . Несмотря на то, что параметр z не является ни признаком объекта, ни состоянием, которое следует распознать, его нельзя исключить из рассмотрения, потому что результат наблюдения x зависит не только от распознаваемого состояния k , но и от скрытого параметра z . Скрытый параметр z будем называть *мешающим* или вмешательством, чтобы отличить его от наблюдения и состояния.

В модели, пример которой мы сейчас рассмотрели, можно говорить о штрафах $W(k, d)$ и об априорных вероятностях $p_K(k)$ скрытых состояний, в данном случае, о вероятностях букв. Однако в данном приложении не определены условные вероятности $p_{X|K}(x | k)$. Причина в том, что вероятность того или иного изображения x зависит не только от буквы, которую это изображение представляет, но и от неслучайного вмешательства, т. е. от то-

го, кто эту букву написал. Здесь мы можем говорить только об условных вероятностях $p_{X|K,Z}(x|k,z)$, т. е. о том, как может выглядеть та или иная буква, написанная тем или иным человеком. Если бы вмешательство было случайно и были бы известны вероятности $p_Z(z)$ для каждого значения z , то можно было бы говорить и о вероятностях $p_{X|K}(x|k)$, потому что их можно было бы вычислить по формуле полной вероятности

$$p_{X|K}(x|k) = \sum_{z=1}^3 p_Z(z) p_{X|K,Z}(x|k,z).$$

Но условия применения алгоритма не дают основания для предположений о том, как часто тот или иной человек будет писать буквы, которые нужно будет распознавать. Например, не исключается, что в течение всего периода эксплуатации алгоритма распознавания придется распознавать тексты, написанные только одним человеком, но неизвестно каким. При таких статистически недоопределенных условиях требуется разработать стратегию, которая обеспечит требуемое качество распознавания букв не зависимо от того, кто написал эту букву. Задача построения этой стратегии должна быть сформулирована так, чтобы понятие априорные вероятности $p_Z(z)$ вмешательства z не использовалось вообще, так как вмешательство не случайно и такое свойство как вероятность для него не определено.

Мы приведем формулировки наиболее известных небайесовских задач и их решения и укажем некоторые модификации известных задач. Затем покажем, что, несмотря на большое разнообразие небайесовских задач, они обладают общими чертами, которые позволяют решать их с помощью единого конструктивного механизма. И, наконец, мы покажем, что класс небайесовских задач имеет общие черты с классом байесовских задач. А именно, класс стратегий решения небайесовских задач почти полностью совпадает с классом байесовских стратегий. Любая небайесовская стратегия, как и байесовская, реализуется в виде разбиения пространства вероятностей на выпуклые конусы.

Лекция 3

Две статистические модели распознаваемого объекта

Условное распределение вероятностей $p_{X|K}: X \times K \rightarrow \mathbb{R}$ наблюдений $x \in X$ при условии, что объект находится в состоянии $k \in K$ — это центральное понятие, на котором основаны различные задачи распознавания. Сейчас самое подходящее время, чтобы привести простейшие примеры таких распределений, на которых будем иллюстрировать как предыдущие, так и последующие теоретические построения. В этой лекции мы прервем на время последовательность изложения курса, чтобы описать две простейшие функции $p_{X|K}$, наиболее употребляемые в качестве моделей распознаваемого объекта.

Лекция 4

Обучение распознаванию образов

4.1. Мифы об обучении в распознавании

Становление любой области науки и техники, которая существенно изменяет человеческую деятельность, проходит, как правило, через следующие три этапа.

На первом этапе сочиняются сказки о том или ином чудесном устройстве, с помощью которого станет возможным то, что до сих пор было невозможно. Скажем, это мечта о ковре-самолете, который позволит летать по воздуху. На втором этапе конструируются модели, похожие скорее на игрушки, чем на что-то серьезное. При определенной доброжелательности можно считать, что эти модели уже делают сказку былью (т. е. как-то уже летают), хотя они и непригодны для практического использования. Только на третьем этапе появляется изделие (в рассматриваемом примере это самолет), которое поначалу мало, а затем все больше и больше становится пригодным для практического использования.

Нет сомнений в значимости третьего этапа. Однако очень важно понимать необходимость и первых двух этапов. Именно при сочинении сказок формируется цель исследования и разработки, причем в ее чисто прикладном, потребительском аспекте. При конструировании игрушечных моделей проверяется принципиальная осуществимость замысла, проверяется, не противоречит ли замысел фундаментальным законам природы, например закону о сохранении энергии. Именно при конструировании игрушек достигаются основные теоретические результаты, которые лягут в основу будущего изделия. При конструировании изделия уже не до теории, так как на повестку дня выносятся проблемы совсем другого порядка. Поэтому не следует торопливо проскакать через сочинение сказок и конструирование игрушек даже при очень нетерпеливом стремлении к скорому практически ощутимому результату. Излишняя поспешность может привести к глубоким и длительным отрицательным последствиям. Не следует

пренебрежительно относиться ни к сочинению мифов, ни к конструированию игрушек, ни к разработке практического изделия. Но очень желательно как можно четче различать эти этапы: не называть прекрасные мечты теоретическим исследованием, а идейно богатую модель — изделием.

Современное распознавание является удивительной смесью мифов, игрушек и изделий, и в наибольшей степени это относится к той части проблемы, которая называется обучением распознаванию образов.

В предыдущих лекциях мы уже видели, что для конструирования стратегии распознавания, т. е. функции $q: X \rightarrow K$, требуются знания. Речь идет как о знаниях об объекте распознавания, которые следует вложить в распознающую систему, так и о профессиональных знаниях о распознавании образов. Получение и тех, и других знаний — это длительный, утомительный и зачастую дорогостоящий процесс. Как реакция на эту действительность, возникают мечты о каком-то чудодейственном средстве типа скатерти-самобранки, с помощью которого можно было бы избежать подобных неприятностей.

В распознавании образов эти мечты имеют обычно следующую форму. «Существует система (генетическая, эволюционная, нейронная или с иным экзотическим названием), которая работает в следующем режиме. Сначала система обучается на основе последовательности x_1, x_2, \dots, x_l примеров наблюдений, подаваемых на ее вход. Каждое наблюдение x_i из этой последовательности сопровождается информацией k_i о реакции распознающей системы, которая считается правильной. После наблюдения l примеров такого правильного поведения обучаемая распознающая система оказывается способной давать правильную реакцию k на любое наблюдение x , в том числе и на те, которые не участвовали при обучении. Поскольку информация о требуемой стратегии вводилась не в явном виде, а только через обучающую последовательность, то обучаемую распознающую систему можно научить решать любую задачу распознавания».

Обычно бесполезно добиваться у автора подобных формулировок более конкретного изложения алгоритма обучения и, тем более, формулировки задачи, которая этим алгоритмом решается. Ожидаемые результаты представляются такими замечательными и легко достижимыми, что просто не хочется тратить время на такие мелочи, как однозначно понимаемая формулировка задачи и обоснованный вывод алгоритма для ее решения. Придуманная сказка настолько прекрасна, что любые приземленные вопросы ее только портят.

Реалистический взгляд на проблемы обучения приводит к известным в настоящее время формулировкам задач, которые даны в следующем разделе.

Лекция 5

Линейные дискриминантные функции

5.1. Вводные замечания о линейных дискриминантных функциях

В предыдущих лекциях мы неоднократно обращали внимание на то, что линейные дискриминантные функции заслуживают особого внимания. Во-первых, известно, что для некоторых статистических моделей решение байесовских или небайесовских задач достигается именно с помощью линейных дискриминантных функций.

Во-вторых, некоторые нелинейные дискриминантные функции можно выразить как линейные в так называемом спрямляющем пространстве, как это было показано в разделе ???. Это возможно, когда нелинейная дискриминантная функция $f: X \rightarrow \mathbb{R}$ имеет вид

$$f(x) = \sum_{j \in J} \alpha_j f_j(x)$$

при известных функциях $f_j: X \rightarrow \mathbb{R}$, $j \in J$, и неизвестных коэффициентах α_j . В этом случае поиск дискриминантной функции, которая нелинейна в исходном пространстве X , сводится к поиску линейной дискриминантной функции в $|J|$ -мерном спрямляющем пространстве Y . Пространство X отображается в пространство Y так, что точке $x \in X$ ставится в соответствие точка $y \in Y$, j -я координата которой равна $f_j(x)$.

В-третьих, важный теоретический результат, указанный в первой лекции, состоит в том, что в пространстве вероятностей любая байесовская стратегия реализуется линейными дискриминантными функциями, и это справедливо для любой статистической модели распознаваемого объекта.

В-четвертых, известно, что емкость класса стратегий, реализуемых линейными дискриминантными функциями в n -мерном пространстве, равна $n + 2$ и поэтому задача обучения корректна. Это значит, что стратегия, построенная по конечному обучающему множеству достаточно большой длины, не слишком отличается от стратегии, найденной по

статистической модели. Более того, в этом случае взаимосвязь точности, надежности и длительности обучения выражается простой и удобной для практического применения формулой.

Все эти преимущества не имели бы большого значения при отсутствии хорошо разработанных методов построения линейных дискриминантных функций, которые составляют содержание данной лекции. Мы увидим, что эти методы как бы взаимодействуют друг с другом, действуют сообща так, что понимание свойств и алгоритмов решения одной задачи позволяет обнаружить неожиданные свойства и ранее неизвестные алгоритмы решения другой задачи.

Лекция 6

Самообучение

Если к правде святой
Мир дорогу найти не умеет —
Честь безумцу, который навеет
Человечеству сон золотой.

По безбрежным блуждая просторам,
Нам безумец открыл Новый Свет;
Нам безумец дал Новый Завет —
Ибо этот безумец был Богом.
Если б завтра Земли нашей путь
Осветить наше Солнце забыло —
Завтра целый бы мир осветила
Мысль безумца какого-нибудь.

Пьер-Жан Беранже, «*Безумцы*», 1833, перевод В.Курочкина.

6.1. Вводные замечания об особом характере лекции

Эта лекция посвящена самообучению в распознавании образов, известному также как обучение без учителя, неконтролируемое обучение, а также под другими названиями. Содержание этой лекции мы изложим в несколько иной форме, чем в предыдущих лекциях. Наши предыдущие рассказы были основаны на понятиях, сформировавшихся вне распознавания и уже в своей завершённой форме перешедших в распознавание. Поэтому исходным моментом предыдущих лекций являлась однозначно понимаемая задача, за которой следовал ее анализ и алгоритмы решения. При таком построении лекции оставался за кадром длительный период жизни научной идеи, предшествующий ясно осознанной формулировке задачи. Этот этап научного исследования, как правило, имеет яркую эмоциональную окраску и наполнен драматизмом.

В отличие от фундаментальных понятий, перешедших в распознавание из других областей, идея самообучения родилась внутри самого распознавания. Она еще не достигла такого почтенного возраста, как байесовские или небайесовские подходы. Однако в молодом возрасте самообучения есть что-то очень ценное: весь период его существования все еще можно окинуть одним взглядом и увидеть все те рытвины и колдобины, по которым оно должно было пройти, прежде чем воплотиться в четкой постановке задачи и ее решении.

Драматические обстоятельства рождения новых знаний представляют собой столь привлекательную тему для обсуждения, что мы с трудом удерживаемся от соблазна показать на ряде примеров, как во все времена и на всех континентах научные открытия постигали очень сходные судьбы. Не совсем серьезно, но и не просто в шутку, можно сказать, что открытие Америки как бы вобрало в себя все то общее, что присуще и научным открытиям.

- Открытие Америки Колумбом стало возможным благодаря мощной финансовой поддержке проекта.
- Колумб получил заказ на работу только после того, как сформулировал ожидаемый результат, имеющий явную прикладную направленность, понятную заказчикам,— открытие нового пути из Европы в Индию.
- Уверенность Колумба, что он решит важную практическую задачу, основывалась на ложных предпосылках, и ожидаемого результата он не достиг. Наоборот, в процессе выполнения проекта обнаружилось существенное препятствие — целый континент, что убедительно доказывало невозможность достижения поставленной цели.
- Несмотря на явно отрицательные результаты, Колумбу удалось убедить заказчиков, что проект успешно выполнен в полном соответствии с заданием.
- После выполнения проекта никто, включая и исполнителя, не заметил того действительно важного положительного результата, который все же был достигнут. Более того, если бы кто-то и обратил на него внимание, он едва ли осознал действительное его значение: никого не интересовали новые, до сих пор неизвестные территории, а лишь новые пути к уже известным территориям.
- Когда возникла практическая потребность в знаниях, фактически полученных Колумбом, о его результатах уже никто не помнил и пришлось начинать новый проект с новыми заказчиками и новым финансированием.
- Континент, открытый Колумбом, назван вовсе не его именем, а именем совсем другого человека, который лишь повторил то, что сделал Колумб намного раньше.
- Сам Колумб до конца своих дней так и не понял, что же он на самом деле открыл, и был уверен (а может, только делал вид?), что открыл новый путь в Индию.
- Сейчас уже всем известно, что Америку открыл вовсе не Колумб. Задолго до Колумба о существовании Америки знали скандинавские народы и практически использовали это знание.

Форма данной лекции обусловлена не только ее содержанием, но и нашим желанием показать, насколько замысловаты пути к новым знаниям, пусть даже почти незаметным.

Лекция 7

Введение в структурное распознавание

7.1. Общность статистического распознавания

В предыдущих лекциях мы неоднократно подчеркивали исключительную общность статистического распознавания. В силу этой общности выводы статистической теории распознавания имеют характер законов. Они должны учитываться при решении любой прикладной задачи, будь то диагноз человеческого сердца по электрокардиограмме, или оценка токарного инструмента по звуковому сигналу, который он издает во время работы, или обработка микроизображений кровяных телец, или исследование естественных ресурсов Земли по космическим снимкам. Сейчас, впрочем, как и раньше, мы хотим подчеркнуть не столько огромный спектр возможных приложений статистической теории распознавания, сколько богатство ее формальных свойств. Множества X и K значений наблюдаемых и скрытых параметров могут быть весьма разнообразны в чисто математическом смысле этого слова. Мы скажем, пока еще очень неточно, что эти множества разнообразны по своей структуре.

Когда мы говорим, что такой параметр, как, например, вес объекта, принимает значения из хорошо структурированного множества, мы имеем в виду, что элементы этого множества можно складывать, умножать, сравнивать между собой, и каждая из этих операций имеет смысл в терминах исходной прикладной задачи. Оценки же, которые получает студент на экзамене, принадлежат множеству с более слабой структурой. Это — множество $\{2, 3, 4, 5\}$ (по крайней мере, в некоторых странах). Его элементы можно сравнивать между собой, т. е. говорить, какая оценка лучше, а какая хуже. Но их нельзя складывать или умножать, так как два студента, получивших на экзамене двойку и тройку, ни в каком смысле не эквивалентны

одному студенту, получившему пятерку. На множестве оценок определены отношения $=$, $<$, и $>$, а другие операции не определены, так как им ничего не соответствует в исходной прикладной задаче. Говоря иными словами, множество оценок является полностью упорядоченным множеством и ничем иным. Еще более слабую структуру имеет множество номеров трамвайных маршрутов. Хотя эти номера и выражаются числами, структура множества маршрутов отличается от структуры множества чисел. В отличие от чисел, трамвай №12 никаким образом не заменяют два трамвая №6. В отличие от элементов упорядоченного множества, маршрут №2 ничуть не лучше маршрута №3 и не хуже маршрута №1.

Сейчас мы говорим об общности статистической теории распознавания, имея в виду именно разнообразие формальных свойств множеств X и K , а не разнообразие их прикладных интерпретаций.

Как правило, в распознавании образов наблюдение x состоит не из одного, а из нескольких измерений x_1, x_2, \dots, x_n . В этом случае можно говорить не только о структуре множества значений каждого отдельного признака, но и о структуре множества признаков. Множества признаков (именно признаков, а не их значений!) в различных задачах имеют различную структуру. Поясним это различие на следующих двух примерах.

В первом примере признаки x_1, x_2, \dots, x_n — это записи в истории болезни при первой регистрации пациента у врача. Они касаются температуры тела, давления крови, возраста, частоты пульса, пола, скажем, всего n признаков. Во втором примере речь идет о значениях какого-то одного признака, например, температуры, измеренных n раз через равные интервалы времени после определенного врачебного воздействия. Результатом такого наблюдения является опять-таки совокупность индексированных величин x_i , где индекс i принимает значения на множестве $\{1, 2, \dots, n\}$, как и в первом примере. Хотелось бы, однако, чтобы было очевидным, что структуры множеств индексов в приведенных двух примерах совершенно различные. Так, в первом примере несущественно, является ли третьим признаком именно возраст пациента. В задаче равным счетом ничего не изменится при произвольной перенумерации признаков. Множество признаков в первой задаче лишено какой-либо структуры. Номера $1, 2, \dots, n$ признаков — это не числа, а просто символы в определенном абстрактном алфавите.

Ситуация во втором примере в корне иная. Здесь индекс признака понимается именно, как целое число. Совокупность измеренных значений температуры образует последовательность, и нумерация элементов в этой последовательности не может быть произвольной. Множество индексов здесь обладает ясно видимой структурой, которая отсутствовала в первом примере. Мы еще не раз вернемся к этим вопросам. Сейчас же мы бросили только беглый взгляд на огромное разнообразие, которое кроется за словами «Пусть множество наблюдений X и множество состояний K — два конечных множества», которые рефреном повторялись в предыдущих лекциях, как исходный пункт для формальных рассуждений.

Общность полученных в предыдущих лекциях результатов состоит в

том, что они справедливы для множеств X и K произвольной структуры. Отсюда, конечно же, не следует, что для решения каждой конкретной задачи достаточно знать лишь самые общие методы статистического распознавания. В каждом конкретном случае задачу следует формулировать как можно более конкретно, т. е. наделять множества X и K структурами, которые наиболее полно отражают ее прикладное содержание. Затем, уже для определенного частного случая, следует исследовать, во что выливаются общие рекомендации. Это нелегкий путь, и структурное распознавание, о котором пойдет речь в последующих лекциях, содержит рекомендации, как его проходить.

К счастью, некоторые приложения можно выразить в рамках формализма, который хорошо исследован в классической математической статистике в ее наиболее развитой части — статистике случайных чисел. Ее рекомендации основаны на таких понятиях, как математическое ожидание, дисперсия, корреляционная матрица, ковариационная матрица. Все эти понятия определены только для случайных объектов, представленных числами. Существует однако масса приложений, в которых результат наблюдения является не числом, а элементом какого-то иного множества, структура которого отличается от структуры множества чисел. К таким приложениям рекомендации численной статистики непосредственно не применимы. Стремление во что бы то ни стало решать эти задачи методами статистики случайных чисел приводит к недопустимой деформации исходной прикладной задачи. В конечном итоге разработанный алгоритм решает совсем не ту задачу, которую следовало решить.

В этом смысле в наименее благоприятном положении оказались задачи распознавания изображений. Мы имеем в виду не абстрактные изображения, а конкретно результат наблюдения объекта с помощью, скажем, телевизионной камеры. Изображение — это исключительно своеобразный объект машинного анализа. Множества изображений, представляющие интерес в том или ином приложении, не принадлежат к классу множеств, досконально изученных в процессе многовековых математических исследований. Это не выпуклые множества, не подпространства, не эллипсы, не шары, и вообще, это не то, что хорошо известно. Это совсем иные множества.

Существенные положительные результаты по распознаванию изображений не могут быть получены на основании одних лишь общих рекомендаций статистической теории распознавания. Необходимо изучать и учитывать исключительную специфичность изображения, как объекта формального анализа.

Лекция 8

Распознавание марковских последовательностей

8.1. Вводные замечания о последовательностях

Последовательность — это, пожалуй, простейшая структура, с которой уместно начать изложение основных понятий структурного распознавания. Даже на таком простейшем примере, как распознавание последовательностей, можно достаточно выразительно показать, как следует распознавать сложный объект, состоящий из большого количества взаимозависимых частей, и как знание этой зависимости позволяет улучшить распознавание объекта в целом и каждой его части. Мы увидим, что хотя алгоритмы распознавания таких объектов не всегда тривиальны, их реализация не приводит к каким-либо непреодолимым вычислительным трудностям. Мы увидим также, что статистические задачи распознавания, обучения и самообучения, сформулированные ранее в общем виде, в исследуемом частном случае решаются точно и без каких-либо дополнительных упрощающих предположений.

Конечно же, последовательности интересны не только тем, что они — прекрасный объект для формального исследования, но и тем, что они хорошо приспособлены для обширного слоя прикладных задач. Мы приведем два примера таких приложений. При этом ни в коем случае не будем утверждать, что эти прикладные задачи полностью укладываются в рамки формальных методов, которые излагаются в данной лекции. Эти два примера приводятся лишь как иллюстрация абстрактных понятий, в терминах которых эти методы будут изложены.

Пример 8.1 Распознавание текстовой строки. Пусть x_1, x_2, \dots, x_n — строка текста, т. е. последовательность изображений букв. Каждому изображению в этой последовательности соответствует наименование буквы из алфавита K . Цель распознавания состоит в том, чтобы на основании наблюдаемой последовательности x_1, x_2, \dots, x_n изображений сказать нечто убедительное о последовательности k_1, k_2, \dots, k_n букв, представленных на этих изображениях.

Если бы буквы k_i , $i = 1, \dots, n$, были взаимно независимы, решение о последовательности k_1, k_2, \dots, k_n свелось бы к n независимым решениям о каждой букве отдельно. Решение о букве на i -м месте принималось бы на основании одного лишь изображения x_i в последовательности x_1, x_2, \dots, x_n . Такое решение далеко не лучшее в реальной ситуации, когда буквы в строке зависят друг от друга. Эта зависимость (контекст) настолько сильная, что часто, на основании знания букв $k_1, k_2, \dots, k_{i-1}, k_{i+1}, \dots, k_n$ можно уверенно судить о букве k_i , даже когда изображение x_i очень сильно искажено или вообще не известно. Проблема заключается в том, что с самого начала не известна ни одна буква в последовательности k_1, k_2, \dots, k_n и решение о букве k_i нужно принять, когда другие буквы тоже не известны. Следовательно, на основании всей последовательности x_1, x_2, \dots, x_n наблюдаемых изображений необходимо принять решение сразу о всей последовательности k_1, k_2, \dots, k_n букв, учитывая при этом априорные знания о их взаимной зависимости. ▲

Пример 8.2 Медицинская диагностика. Пусть x_1, x_2, \dots, x_n — результат измерения определенного параметра пациента в последовательные моменты времени, а k_1, k_2, \dots, k_n — последовательность его состояний в эти моменты. Эти состояния недоступны для непосредственного наблюдения. Цель распознавания состоит в принятии как можно более точного решения о состоянии пациента в тот или иной момент времени. Естественно, состояния в различные моменты времени зависят друг от друга, и знание этой зависимости позволяет улучшить качество распознавания. ▲

К настоящему времени достаточно глубоко исследованы алгоритмы распознавания последовательностей при определенной формулировке задачи распознавания. Характеристической чертой этих алгоритмов является применение динамического программирования для нахождения оптимальной последовательности k_1, k_2, \dots, k_n скрытых состояний. Первоисточником этих методов являются статьи Ковалевского [?], разработавшего их для распознавания текстовых строк. В работе Винцока [?] этот подход впервые применен для распознавания речевых сигналов. С тех пор эти методы быстро распространились по всему миру, и сейчас они пользуются такой популярностью, которая начинает им уже вредить. Мало-помалу были забыты формулировки задач, которые уместно решать именно с помощью динамического программирования, и оно стало применяться без разбору и в тех приложениях, для которых уместны иные постановки, а следовательно, и иные решения.

В этой лекции мы определим класс статистических моделей распознаваемого объекта и для этого класса сформулируем различные байесовские задачи распознавания. Мы увидим, что одни задачи естественно решаются хорошо известным методом динамического программирования. Для других задач уместны менее известные, но тоже очень хорошие методы.

Лекция 9

Регулярные языки и задачи распознавания

9.1. Регулярные языки

Лекция 10

Двумерные КС-языки и задачи распознавания

10.1. Вводные замечания

Время от времени научная терминология словно издевается над доверчивым читателем и намеренно сбивает его с толку. Это происходит, когда то или иное точное понятие обозначается словом, которое используется и в повседневной жизни, но для не очень точных, размытых понятий. Так, теория катастроф не имеет ничего общего с тем, что происходит в повседневной жизни, когда говорят о каком-то неожиданном бедствии. Точно так же математическая теория игр совсем не является формализацией игры на футбольном поле или шахматной доске.

Понятие «контекстно-свободный язык» представляет пример такого зловредного понятия. Если не знать его точное определение, то сами слова в его названии могут сильно дезориентировать. Может показаться, что речь идет о языке, в котором слова выстраиваются друг за другом без учета контекста, образуя некоторую абракадабру из взаимно независимых фрагментов. Это, конечно, совсем не так. Понятие контекстно-свободного языка имеет точное математическое определение, которое мы приведем. Однако догадаться до него невозможно именно из-за выразительного названия, которое в обыденной жизни обозначает нечто совсем другое.

В данной лекции описывается формализм для конструктивного задания множеств изображений и распределений вероятностей на этих множествах. На основе этих данных формулируются и решаются задачи распознавания, подобные рассмотренным в предыдущих лекциях 8 и 9. Отличается же данная лекция от двух предыдущих тем, что объекты распознавания — это не обязательно одномерные последовательности. Это могут быть двумерные или многомерные массивы или иметь еще более общую

структуру. Мы увидим, что предложенный формализм является естественным обобщением контекстно-свободных грамматик и языков в иерархии Н.Хомского, которые в свою очередь являются обобщением ранее рассмотренных регулярных грамматик и языков.

Оглавление

Введение	3
Предисловие к русскому изданию	3
Предисловие к английскому изданию	5
Письмо аспиранта Иржи Пехи до опубликования лекций	7
Письмо авторов аспиранту Иржи Пехе	9
Основные понятия и обозначения	11
Лекция 1. Байесовские задачи распознавания	15
1.1. Вводные замечания	15
1.2. Формулировка байесовских задач	16
1.3. Два свойства байесовских стратегий	17
1.4. Два частных случая байесовских задач	22
1.4.1. Вероятность ошибочного решения о состоянии	22
1.4.2. Байесовская стратегия отказа от распознавания	24
1.5. Обсуждение	26
1.6. Библиографические заметки	40
Лекция 2. Небайесовские задачи распознавания	41
2.1. Жесткая ограниченность байесовского подхода	41
2.1.1. Функция потерь	42
2.1.2. Априорные вероятности состояний	42
2.1.3. Условные вероятности наблюдений	44
2.2. Формулировка известных и новых небайесовских задач	45
2.2.1. Задача Неймана–Пирсона	45
2.2.2. Обобщенная задача Неймана–Пирсона для двух опасных состояний	48
2.2.3. Минимаксная задача	48
2.2.4. Задача Вальда	49

2.2.5. Статистические решения при неслучайных воздействиях	51
2.3. Двойственные задачи линейного программирования	52
2.4. Решения небайесовских задач на основе теорем двойственности	58
2.4.1. Решение задачи Неймана–Пирсона	59
2.4.2. Решение обобщенной задачи Неймана–Пирсона для случая двух опасных состояний	62
2.4.3. Решение минимаксной задачи	64
2.4.4. Решение задачи Вальда для случая двух состояний	66
2.4.5. Решение задач Вальда для произвольного количества состояний	69
2.4.6. Испытание сложных случайных гипотез	71
2.4.7. Испытание сложных неслучайных гипотез	72
2.5. Заключительные замечания	73
2.6. Обсуждение	73
2.7. Библиографические заметки	91

Лекция 3. Две статистические модели распознаваемого объекта 93

3.1. Условная независимость признаков	93
3.2. Гауссово распределение вероятностей	95
3.3. Обсуждение	98
3.4. Библиографические замечания	121

Лекция 4. Обучение распознаванию образов 123

4.1. Мифы об обучении в распознавании	123
4.2. Три формулировки задач обучения в распознавании	124
4.2.1. Обучение как максимально правдоподобное оценивание	127
4.2.2. Обучение по неслучайному обучающему множеству	128
4.2.3. Обучение по минимизации эмпирического риска	129
4.3. Основные понятия статистической теории обучения	130
4.3.1. Неформальное описание проблем обучения распознаванию	131
4.3.2. Основы статистической теории обучения распознаванию по Вапнику и Червоненкису	135
4.4. Критический взгляд на статистическую теорию обучения	143
4.5. набросок детерминированного обучения	145
4.6. Обсуждение	150
4.7. Библиографические заметки	159

Лекция 5. Линейные дискриминантные функции 161

5.1. Вводные замечания о линейных дискриминантных функциях	161
5.2. Путеводитель по лекции	162
5.3. Задачи Андерсона	165
5.3.1. Эквивалентные формулировки задач Андерсона	165
5.3.2. Неформальный анализ задачи Андерсона	167

5.3.3. Определение вспомогательных понятий для задачи Андерсона	170
5.3.4. Решение исходной задачи Андерсона	172
5.3.5. Формальный анализ обобщенной задачи Андерсона . . .	174
5.4. Линейное разделение конечных множеств точек	183
5.4.1. Формулировка задачи и ее анализ	183
5.4.2. Алгоритмы линейного разделения конечных множеств точек	186
5.4.3. Алгоритмы ϵ -оптимального разделения конечных множеств точек	191
5.4.4. Построение фишеровских классификаторов	193
5.4.5. Другие модификации алгоритмов Козинца	195
5.5. Решение обобщенной задачи Андерсона	199
5.5.1. ϵ -решение задач Андерсона	199
5.5.2. Линейное разделение бесконечных множеств точек . . .	203
5.6. Обсуждение	206
5.7. Связь с toolbox'ом	227
5.8. Библиографические заметки	228

Лекция 6. Самообучение 229

6.1. Вводные замечания об особом характере лекции	229
6.2. Предварительное неформальное определение самообучения .	231
6.3. Самообучение в перцептроне	232
6.4. Эмпирический байесовский подход Роббинса	240
6.5. Квадратичная кластеризация	247
6.6. Алгоритмы самообучения и их анализ	252
6.6.1. Задача распознавания	252
6.6.2. Задача обучения	252
6.6.3. Задача самообучения	254
6.6.4. Алгоритм самообучения	255
6.6.5. Анализ алгоритмов самообучения	256
6.6.6. Алгоритм решения задачи Роббинса и ее анализ	266
6.7. Обсуждение	269
6.8. Связь с TOOLBOX	289
6.9. Библиографические заметки	289

Лекция 7. Введение в структурное распознавание 291

7.1. Общность статистического распознавания	291
7.2. Структурное распознавание изображений	293
7.2.1. Множество наблюдений	293
7.2.2. Множество скрытых параметров изображений	296
7.2.3. Роль обучения в распознавании изображений	298
7.3. Основные понятия структурного распознавания	301
7.4. Обсуждение	306
7.5. Библиографические замечания	323

Лекция 8. Распознавание марковских последовательностей	325
8.1. Вводные замечания о последовательностях	325
8.2. Марковская модель распознаваемого объекта	327
8.3. Распознавание стохастического автомата	331
8.3.1. Распознавание стохастического автомата; формулировка задачи	331
8.3.2. Алгоритм распознавания стохастического автомата . .	332
8.3.3. Матричное представление вычислительной процедуры	333
8.3.4. Статистическая интерпретация матричных произведений	334
8.3.5. Распознавание марковских объектов при неполных данных	337
8.4. Наиболее вероятная последовательность скрытых параметров	339
8.4.1. Различие между распознаванием объекта в целом и распознаванием его составных частей	339
8.4.2. Формулировка задачи поиска наиболее вероятной последовательности состояний	340
8.4.3. Сведение задачи к поиску кратчайшего пути на графе	340
8.4.4. Поиск кратчайшего пути на графе	342
8.4.5. О формальном анализе задач	344
8.4.6. Обобщенные матрицы и их умножение	347
8.4.7. Распознавание подпоследовательности состояний	350
8.5. Поиск последовательности наиболее вероятных состояний . .	353
8.6. Марковские объекты с ациклической структурой	359
8.6.1. Статистическая модель объекта	359
8.6.2. Вычисление вероятности наблюдения	361
8.6.3. Наиболее вероятная совокупность скрытых параметров	364
8.7. Формулировка задач обучения и самообучения	365
8.7.1. Наиболее правдоподобное оценивание модели в режиме обучения	367
8.7.2. Минимаксная оценка модели	367
8.7.3. Настройка алгоритма распознавания	368
8.7.4. Задача самообучения	369
8.8. Построение наиболее правдоподобной модели в режиме обучения	369
8.9. Минимаксное оценивание марковской модели	375
8.9.1. Формулировка алгоритма и его свойств	375
8.9.2. Анализ задачи минимаксного оценивания	379
8.9.3. Минимаксное оценивание марковской модели	388
8.10. Настройка алгоритма распознавания последовательностей . .	389
8.11. Наиболее правдоподобное оценивание при самообучении . . .	391
8.12. Обсуждение	395
8.13. Связь с TOOLBOX'ом	420
8.14. Библиографические замечания	420

Лекция 9. Регулярные языки и задачи распознавания	421
9.1. Регулярные языки	421
9.2. Другие способы представления регулярных языков	423
9.2.1. Регулярные языки и автоматы	423
9.2.2. Регулярные языки и грамматики	424
9.2.3. Регулярные языки и регулярные выражения	425
9.2.4. Пример регулярного языка, представленного различными способами	426
9.3. Модификации регулярных языков	428
9.3.1. Размытые автоматы и размытые языки	429
9.3.2. Штрафные автоматы и соответствующие языки	430
9.3.3. Простейшая задача на наилучшее соответствие	431
9.4. Заключение по первой части лекции	433
9.5. Левенштейнова аппроксимация последовательности	435
9.5.1. Предварительная формулировка задачи	435
9.5.2. Функции Левенштейна	435
9.5.3. Известный алгоритм вычисления левенштейнова отклонения	436
9.5.4. Свойства функций Левенштейна	439
9.5.5. Формулировка задачи и ее обсуждение	442
9.5.6. Формулировка основного результата и его обсуждение	443
9.5.7. Обобщенные свертки и их свойства	445
9.5.8. Конволюционная формулировка задачи и основного результата	452
9.5.9. Доказательство основного результата	454
9.5.10. Интерпретация основного результата	465
9.6. Обсуждение	469
9.7. Связь с TOOLBOX'ом	504
9.8. Библиографические замечания	504
 Лекция 10. Двумерные КС-языки и задачи распознавания	 507
10.1. Вводные замечания	507
10.2. Неформальное объяснение двумерных грамматик и языков	508
10.3. Двумерные контекстно-свободные грамматики и языки	513
10.4. Обобщенный алгоритм Кока—Янгера—Касами	515
10.5. Общая структурная конструкция	517
10.5.1. Структурная конструкция для определения множеств распознаваемых объектов	519
10.5.2. Основная задача структурного распознавания	523
10.5.3. Вычислительная процедура решения основной задачи	523
10.6. Обсуждение	527
10.7. Библиографические заметки	533
 Список литературы	 535