

Классификация алкогольной продукции по спектрам флуоресценции с обработкой данных методом машинного обучения

Саакян А. В.¹, Аленичев М. К.², Левин А. Д.^{1,2}

¹ ФГАОУ ВО «Московский физико-технический институт» (национальный исследовательский университет), г. Долгопрудный, Московская область, Россия

² ФГУП «Всероссийский научно-исследовательский институт оптико-физических измерений», г. Москва, Россия, e-mail: levin-ad@vniiofi.ru

Аннотация: Разработан экспресс-метод определения географического происхождения (ГП) коньяков по спектрам эмиссии и синхронного сканирования флуоресценции. Исследовались 43 образца бренди и коньяков, которые по ГП были разделены на три класса – произведенные в регионах России за исключением Дагестана, произведенные в Дагестане и произведенные в Армении. Обработка спектров производилась на основе машинного обучения с построением обучающего и тестового наборов. Правильность классификации для образцов тестового набора

Ключевые слова: спектры флуоресценции, коньяки и коньячные дистилляты, машинное обучение, контроль подлинности, географическое происхождение.

При контроле подлинности алкогольной продукции одной из важных задач является подтверждение ГП. В современных условиях актуальной является разработка экспресс-методов определения ГП, требующих минимальной пробоподготовки и реализуемых на простом и обладают методы оптической молекулярной спектроскопии. В докладе представлены результаты классификации бренди и коньяков по ГП на основе спектров флуоресценции с обработкой данных методами машинного обучения (МО). Преимущества флуоресцентной спектрометрии по сравнению с известными методами определения ГП, такими как газовая и жидкостная хроматография, масс-спектрометрия, заключаются в экспрессности, существенно более простой пробоподготовке, более дешевых измерительных приборах. Спектры флуоресценции оказываются чувствительными к небольшим различиям в химическом составе образцов, обусловленными, в частности, разным ГП коньячных спиртов. В нашем исследовании для распознавания географического происхождения образцов были выбраны 3 вида флуоресцентных спектров: синхронного сканирования при разности длин волн 50 нм, эмиссии при длинах волн возбуждения 250 нм и 280 нм. Пробоподготовка сводилась к разбавлению образцов коньяка дистиллированной водой в 30 раз. Измерения проводились на спектрофотометре-флуориметре СФФ-2 «Флуоран», разработка ФГУП «ВНИИОФИ». Было исследовано 43 образца бренди и коньяков, которые по ГП были разделены на 3 класса – произведенные в различных регионах Российской Федерации, за исключением Республики Дагестан, произведенные в Республике Дагестан и произведенные в Республике Армения. При этом напитки, произведенные в Дагестане и Армении изготавливались исключительно из местных коньячных дистиллятов, а в состав напитков, произведенных в регионах РФ могли входить коньячные дистилляты из Франции и Испании. Был выделен обучающий набор, состоявший из 33 образцов. Для спектров флуоресценции каждого вида строилась модель МО, обучение которой проводилось методом экстремального градиентного бустинга – одного из современных алгоритмов дискриминантного анализа. Для всех моделей имело место правильное на 100 % распознавание при перекрестной валидации. Правильность распознавания образцов тестового набора (состоявшего из всех 43 образцов) составила 97,6% для моделей на основе спектров синхронного сканирования и эмиссии при длине волны возбуждения 250 нм, и 93,0% для модели на основе спектров эмиссии при длине волны возбуждения 280 нм. Разработанная методика и алгоритм машинного обучения могут быть использованы для классификации по географическому происхождению других видов алкогольной продукции, например, вин.