2

Цель работы:

Разработать программный продукт для обработки данных спектров рентгенофлуоресцентных энергодисперсионных спектрометров с целью определения концентраций химических элементов в исследуемых образцах.

Задачи:

* Проанализировать существующие методы обработки данных спектров.
* Разработать алгоритм работы и архитектуру программы для обработки спектров.
* Разработать техническую документацию.
* Реализация программного продукта на языке программирования.
* Подготовка и проведение тестирования программы.

3

Введение

В настоящее время для определения концентраций химических элементов в технологических растворах и материалах широко используются рентгеновские спектрометры. Они обладают рядом преимуществ, по сравнению с методами химического анализа. Тем не менее, при использовании спектрометра требуется привлечение высококвалифицированных специалистов для разработки качественных методик определения концентраций.

Процесс определения химических концентраций при помощи спектрометра можно разбить на три этапа:

1. Получение рентгеновского спектра аппаратной частью спектрометра.
2. Предобработка спектра с целью получения из него набора данных, релевантных данной задаче.
3. Расчет концентраций программным обеспечением с использованием полученного набора данных.

Предметом настоящего исследования является этап 3.

4

Существуют различные методы расчета концентраций, но наиболее точным является метод корректировки по интенсивностям. В основе метода лежит принцип, что концентрация химического элемента пропорциональна интенсивности в линиях его характеристического рентгеновского излучения. Но на практике на интенсивность в линиях определяемого элемента оказывают влияние различного рода излучения от других химических элементов, присутствующих в пробе, а также фоновое излучение от рентгеновской трубки и рассеянное излучение.

5

Формально метод корректировки по интенсивностям можно представить следующим образом:

существует  – множество значений интенсивностей излучения в спектре,

и некоторая функциональная зависимость , связывающая множество значений интенсивностей со значением концентрации. В качестве функции, приближающей эту зависимость, как правило, используется уравнение полиномиального вида

(1)

где С – определяемая концентрация, – свободный член, , – коэффициенты.

Это уравнение называется градуировкой.

Таким образом, для градуировки необходимо определить коэффициенты , , .

6

Постановка задачи

Для получения градуировки проводится измерение спектров от образцов, для которых концентрации определены методом химического анализа, получают значения интенсивности излучений химических элементов и при помощи регрессионного анализа получают искомые коэффициенты.

Так как взаимное влияние линий спектра имеет сложный характер, градуировки для каждого элемента выполняются опытными высококвалифицированными специалистами, что является недостатком этого метода определения концентраций. Для решения этой проблемы предлагается использовать методы машинного обучения.

В качестве модели исследования выбрана нейросетевая модель, так как они эффективны при моделировании сложных нелинейных зависимостей и более точно описывают наборы данных по сравнению с линейными методами статистики.

7

Для решения задачи регрессии выбрана многослойная нейронная сеть с обратным распространением ошибки. Нейронная сеть обратного распространения состоит из нескольких слоев нейронов, причем каждый нейрон слоя i связан с каждым нейроном слоя i+1, т. е. это полносвязная сеть. Каждый нейрон в сети имеет некоторое состояние

где – числовое значение на i-ом входе нейрона, – числовое значение веса на i-ом входе нейрона.

Значение на выходе нейрона определяется как

где - некоторая функция активации.

8

Для того, чтобы определить ошибку нейросети на выборке нужно определить функцию потерь на одном объекте обучающей выборки. Для задач регрессии используется либо абсолютная ошибка

либо квадратичная ошибка

где – целевое значение на выходе, – реальное значение на выходе.

Среднее всех потерь по всем объектам обучающей выборки – эмпирический риск, определяется по формуле

где k-количество объектов в обучающей выборке.

Таким образом, задача обучения нейросети заключается в минимизации эмпирического риска, путем подстройки весов , то есть сводится к оптимизации функции .

Оптимизация может производится, например, методом градиентного спуска, то есть на каждой итерации веса изменяется по формуле

где – параметр, определяющий скорость обучения.

9

Задача настоящего исследования – найти такие параметры сети и алгоритмы оптимизации, при которых задача получения градуировки будет решаться с приемлемым качеством и за разумное время. Качество полученной градуировки проверяется на контрольном массиве спектра и оценивается коэффициентом детерминации:

где – дисперсия остатков, т.е. необъясненная моделью дисперсия, – общая дисперсия.

Для проведения исследований отсняты массивы спектров четырех видов материалов с большим диапазоном химических концентраций.

Исследования и тесты программного продукта будут проводятся с использованием Keras — открытая нейросетевая библиотека, написанная на языке Python.

Возможные темы публикаций для ВКР.

Работа со спектрометром предполагает два основных сценария.

Лабораторный сценарий.

1. Лаборант загружает в спектрометр пробу или набор проб с заранее известными химическими составами и приблизительно известными концентрациями химических элементов.
2. Исходя из знания состава, приблизительных концентраций и набора химических элементов, для которых требуется определить концентрацию, выбирает градуировки из имеющейся базы градуировок. Каждая градуировка позволяет определить концентрацию для одного определенного химического элемента.
3. Запускает измерение спектров и программное обеспечение автоматически определяет требуемые концентрации.

Проблема – лаборант должен знать тип пробы. Если тип пробы определен ошибочно или химические концентрации в пробе значительно отличаются от тех, для которых предназначена градуировка, и лаборант не проконтролировал это по отображению спектра на графике, то концентрация будет определена с ошибкой, превышающей допустимую. Кроме того, лаборант должен иметь соответствующую квалификацию.

Предлагаемое решение – программа производит предварительную классификацию спектров и после этого предлагает релевантные градуировки либо предупреждает, что использование имеющихся градуировок может дать ошибочный результат.

Тема исследования в статье – алгоритмы классификации спектров. Сравнение результатов классификации на имеющихся реальных спектрах.

Разработка градуировки спектрометра.

Градуировка представляет собой программный объект, в котором содержатся:

1. Значения настроек спектрометра, при которых должен быть отснят спектр (экспозиция в секундах, коэффициент усиления детектора и т.д.),
2. Градуировочное уравнение, по которому определяются концентрации.

Градуировочное уравнение имеет, как правило, вид полиномиальной регрессии:

где I – сумма фотонов в линиях характеристического излучения, зафиксированных детектором, k – коэффициенты регрессии.

Разработка градуировки производится методистом в отдельном приложении (Рис. 1) и требует высокой квалификации.

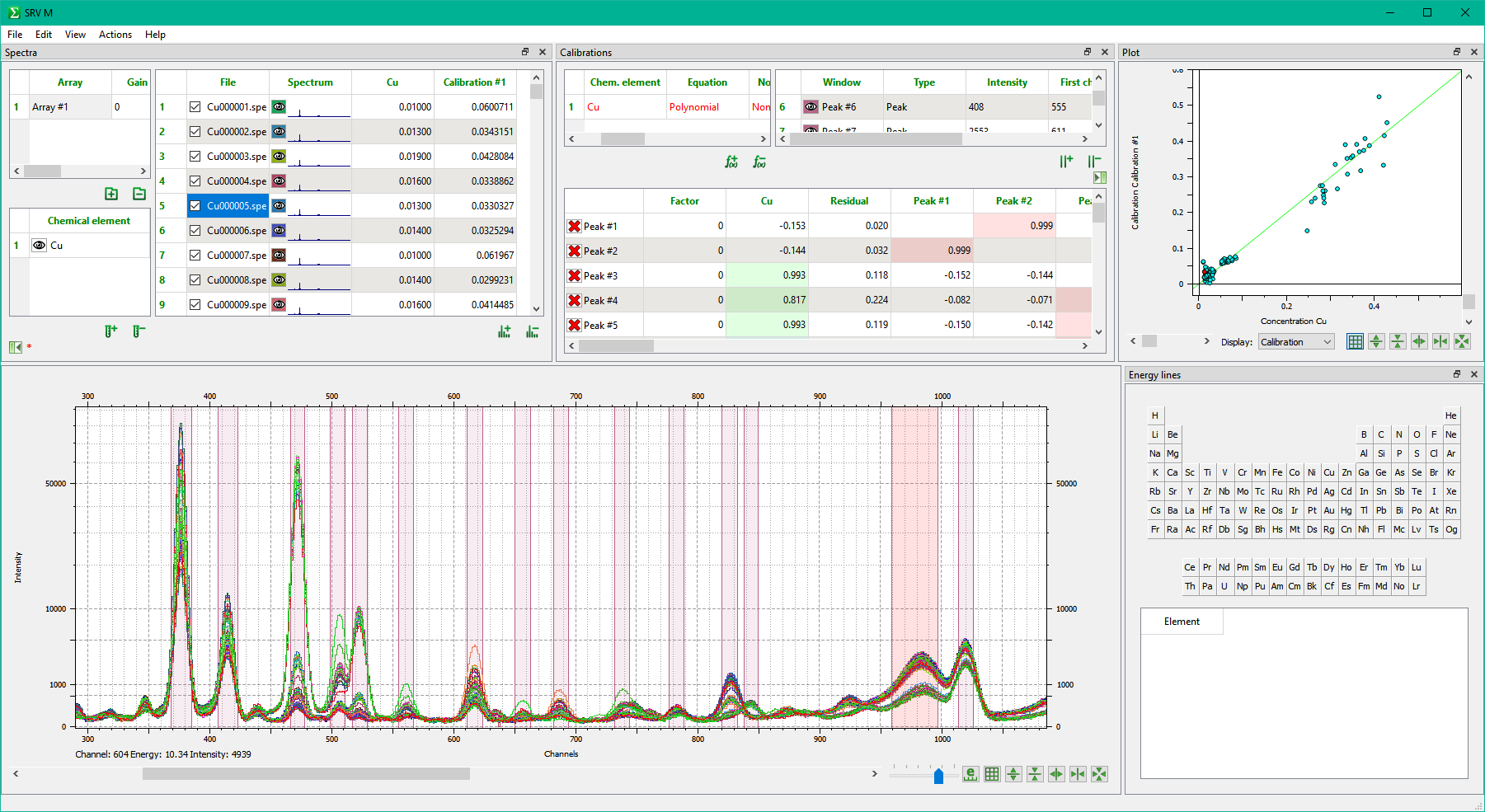


Рис. 1. Приложение SRVM для разработки градуировки.

Работа состоит в следующем:

1. Методист загружает массив отснятых спектров от градуировочных проб. Концентрации химических элементов в этих пробах определены химической лабораторией и используются как референтные значения для разработки градуировки.
2. Из спектров отбираются аналитические и влияющие пики. Если между значениями интенсивностей в пиках имеются корреляции, то есть возможность ввести в градуировку переменную, представляющую произведение интенсивностей или вообще некоторую функцию интенсивностей.
3. Производится расчет коэффициентов регрессии и оценка качества градуировки на проверочном массиве спектров.

Проблема – качество градуировки критически зависит от квалификации и опыта методиста. Как правило, персонал, использующий спектрометр в своей работе такими квалификацией и опытом не обладает, что приводит к необходимости привлечения соответствующих специалистов.

Предлагаемое решение – использовать алгоритмы машинного обучения для построения градуировочного уравнения.

Тема исследования в статье – алгоритмы построения регрессии для использования в градуировках спектрометров.