

МНОГОФАКТОРНЫЙ АНАЛИЗ И ПРОГНОЗ СВОЙСТВ ОРГАНИЧЕСКИХ СОЕДИНЕНИЙ

Ю.В. Передерин

В работе предложен универсальный метод аппроксимации при помощи искусственных нейронных сетей как альтернатива обработки и прогноза свойств различных органических соединений при помощи линейной регрессии

Для проведения комплексного и представительного многофакторного анализа и прогноза свойств органических соединений (ОС) была создана электронная база данных, в которую вошли следующие параметры указанных соединений: формула, название, молекулярная масса, плотность при нормальных условиях (г/см^3), температуры плавления, кипения, вспышки, самовоспламенения, а также верхний и нижний концентрационные пределы воспламенения паров ОС в смеси с воздухом. Все данные для исследований были взяты из справочной литературы [1-5]. В состав базы данных вошли 90 веществ, но в связи с тем, что для некоторых веществ отсутствовали одно или более значений указанных параметров, количество веществ сократилось до 46 за вычетом веществ с неполным набором характеристик.

После внесения изменений, связанных с требованиями к исходным электронным базам программы анализа данных Deductor, была начата сама обработка как таковая.

На первом этапе входные характеристики ОС внутри анализирующей программы были разделены на «входящие» и «выходящие», т.е. на те, на основе которых будет основываться прогнозирование, и на те, прогнозирование которых будет осуществляться созданным в дальнейшем математическим аппаратом, соответственно, а именно: молекулярная масса, плотность при нормальных условиях, температуры плавления, кипения, самовоспламенения, а также верхний и нижний концентрационные пределы воспламенения паров ОС в смеси с воздухом попали в разряд «входных», а температура вспышки – в разряд «выходных» параметров, в то время, как формула и название ОС попали в разряд информационных данных (при проведении данного вида анализа не учитываются лингвистические данные). На этом подготовка данных для начала проведения многофакторного анализа закончена.

Далее внутри программы была проведена попытка обнаружения

коррелирующих данных. В результате были получены следующие данные:

– используя линейную регрессию, температуру вспышки для существующей базы ОС, можно прогнозировать с погрешностью 80,4...100%, не учитывая ни одного параметра, т.е. ,практически, наугад;

– с погрешностью 43,0...80,4% – основываясь на значениях температуры кипения ОС;

– с погрешностью 39,6...43,0% – основываясь на значениях температуры кипения и температуры плавления ОС;

– с погрешностью 34,4...39,6% – основываясь на значениях температуры кипения, температуры плавления с использованием значений верхнего концентрационного предела воспламенения паров ОС;

– с погрешностью 31,3...34,4% – основываясь на значениях температуры кипения, температуры плавления, верхнего концентрационного предела воспламенения паров и температуры самовоспламенения ОС;

– с погрешностью 31,0...31,3% – основываясь на значениях температуры кипения, температуры плавления, верхнего концентрационного предела воспламенения паров, температуры самовоспламенения и плотности ОС при нормальных условиях;

– с погрешностью 29,9...31% – основываясь на значениях температуры кипения, температуры плавления, верхнего концентрационного предела воспламенения паров, температуры самовоспламенения, плотности при нормальных условиях и молекулярной массы ОС;

– с погрешностью 0...29,9% – основываясь на значениях температуры кипения, температуры плавления, верхнего и нижнего концентрационного предела воспламенения паров, температуры самовоспламенения, плотности при нормальных условиях и молекулярной массы ОС, т.е. всех факторов, выбранных как «входные».

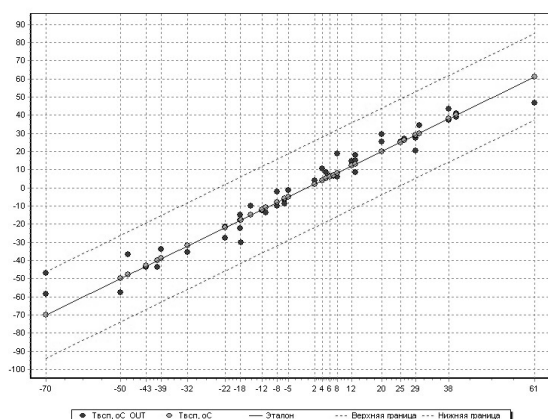


Рисунок. Диаграмма рассеяния температуры вспышки ОС для существующей базы данных при аппроксимации искусственными нейронными сетями с обратным распространением ошибки (пунктирная линия соответствует ошибке аппроксимации 3,3%)

Основываясь на результатах, можно сказать, что использование линейной регрессии позволяет прогнозировать значения температуры вспышки ОС в диапазоне ошибки 0...29,9% аппроксимацией значений созданной базы данных, что может быть приемлемо для некоторых предварительных оценок температуры вспышки, а также выявлено, что основным параметром, определяющим температуру вспышки, является температура кипения ОС.

Использование универсального механизма аппроксимации искусственных нейронных сетей с обратным распространением ошибки дало следующие результаты:

– при аппроксимации значений температуры вспышки при 300 эпохах

обучения и случайном разделении базы на 80% обучающей выборки и 20% тестовой получена математическая модель с диапазоном ошибки прогнозирования 0...3,3% (пунктирная линия на рисунке).

При различных подходах при аппроксимации значений температуры вспышки по созданной базе данных по свойствам ОС было выявлено, что прогнозирование при помощи линейной регрессии производится с ошибкой аппроксимации порядка 0...29,9%, в то время, как при помощи искусственных нейронных сетей ошибка аппроксимации составляет 0...3,3%, что более предпочтительно при проведении фундаментальных и прикладных исследований.

СПИСОК ЛИТЕРАТУРЫ

1. Ротштейн А.П. "Интеллектуальные технологии идентификации" – Электронное издание.
2. Мишулина О.А., Манько С.В. Нейронные сети и устройства нечеткой логики, Приборы и системы. Управление, контроль, диагностика-, 2001- №8- С. 36.
3. Электронный учебник StatSoft – Электронное издание.
4. Гупал А.М., Понамарев А.А., Цветков А.М. Об одном методе индуктивного вывода с подрезанием деревьев решений. // Кибернетика и системный анализ–1993-№5-С. 174.
5. Пожарная опасность веществ и материалов, применяемых в химической промышленности // Справочник – М.: «Химия», 1970 – 336 с.

РОЛЬ ДИСЛОКАЦИОННОЙ СТРУКТУРЫ ВО ВЗРЫВНОМ РАЗЛОЖЕНИИ КРИСТАЛЛОВ АЗИДА СЕРЕБРА

В.И. Крашенинин, Л.В. Кузьмина, Е.Г. Газенаур, Д.В. Добрынин

Показана возможность стимулирования процессов взрывного разложения, иницированного контактным электрическим полем путем контролируемого воздействия на дислокационную структуру кристаллов азидов серебра. Практическая значимость данных исследований определяется возможностью использования полученных экспериментальных данных для целенаправленного изменения взрывной чувствительности по отношению к действию постоянного электрического поля.