

*Исследование средств Data Mining
для интеллектуальной обработки
данных в геологии*



*Дипломная работа
Студента группы ДА-62
СП ИПСА НТУУ «КПИ»
Кирилюка Богдана*



Цель исследования



Предсказание землетрясений с помощью методов интеллектуального анализа данных на основании исторических каталогов.



Сейсмические опасности



Этапы работ



консультация в институте геологии и МЦД

анализ методов кластеризации

сравнение существующих программных средств

исследование данных для обработки

тестирование реализации алгоритма M8



Составные части работы



Инструментарий для исследования

- Программные средства
- Набор методов для анализа
- Получение знаний



Изучение геологической литературы

- Понимание процесса
- Выбор алгоритма
- Проведение анализа

Разработка ПО

- Реализация алгоритма
- Проведение тестирования
- Автоматизация процесса



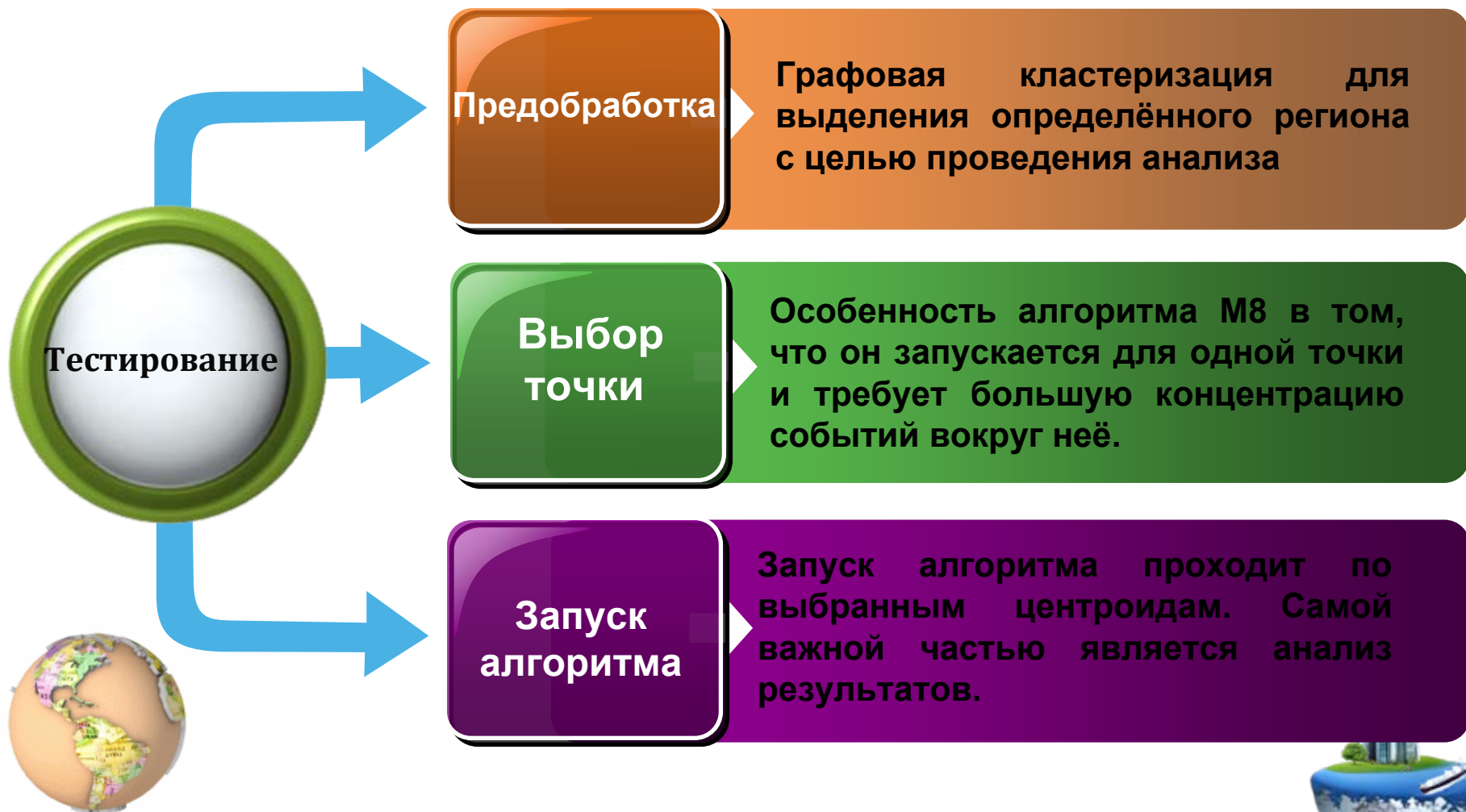
Источники данных



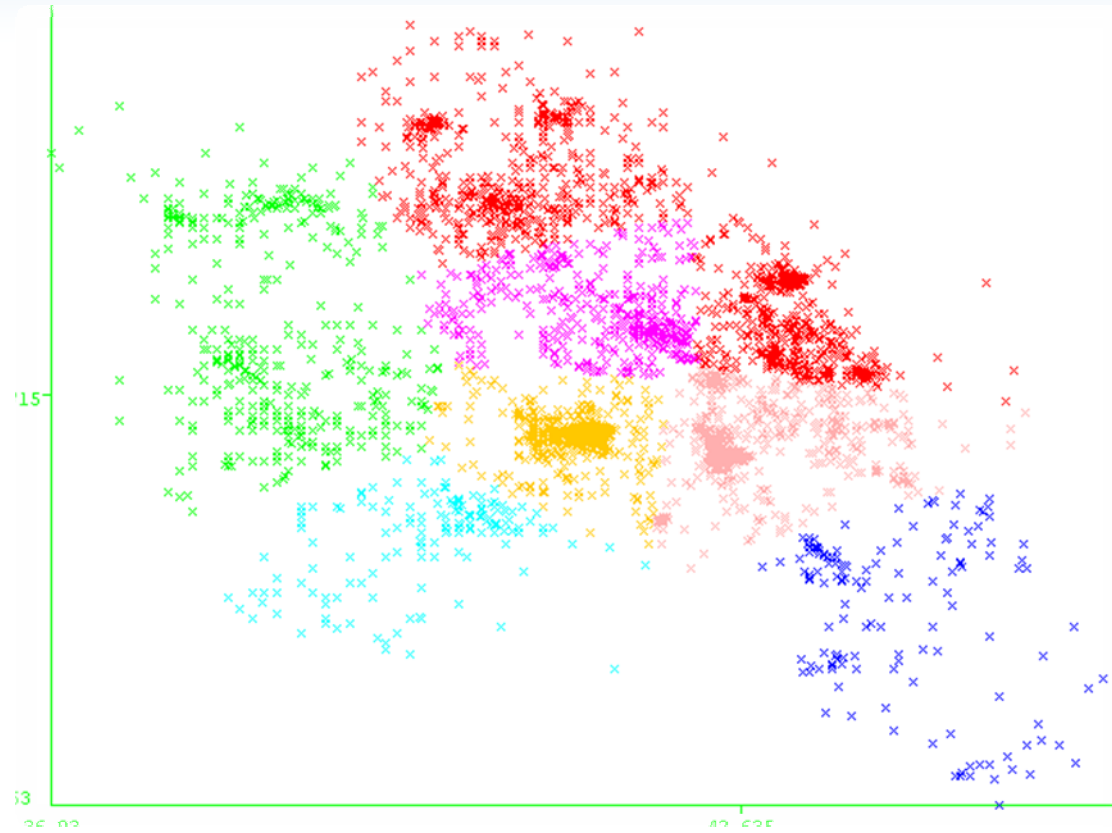
Данные по Украине и Кавказу из официальных источников Мирора и Кенетом для целей
сейсмического анализа за период с 1965 по 2009 год. Данные о
пределах вращательных движений с магнитудой более
3.5 балла.



Тестирование данных



Тестирование данных



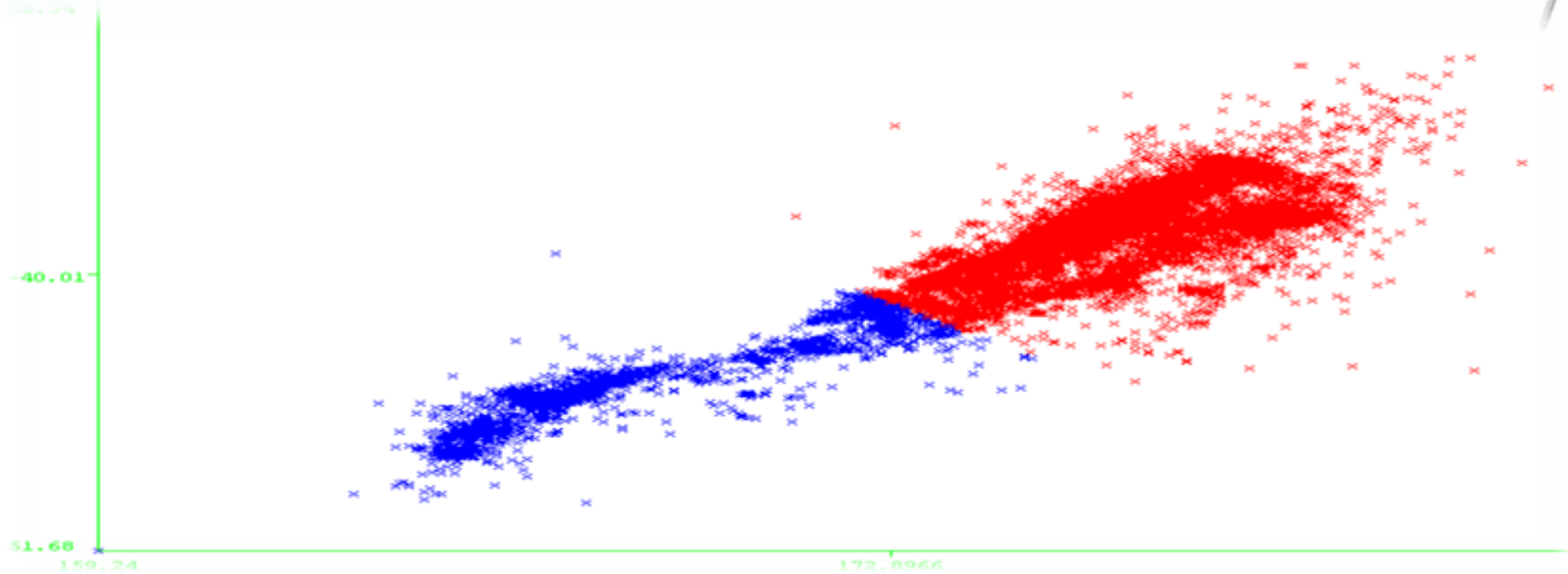
Данные Кавказского региона после кластеризации алгоритмом K-means



При запуске алгоритма M8 на всех центроидах произошло исключение, из-за недостатка данных для расчёта прогноза



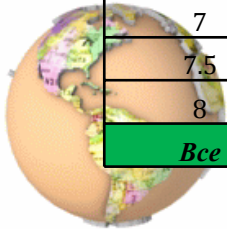
Тестирование данных



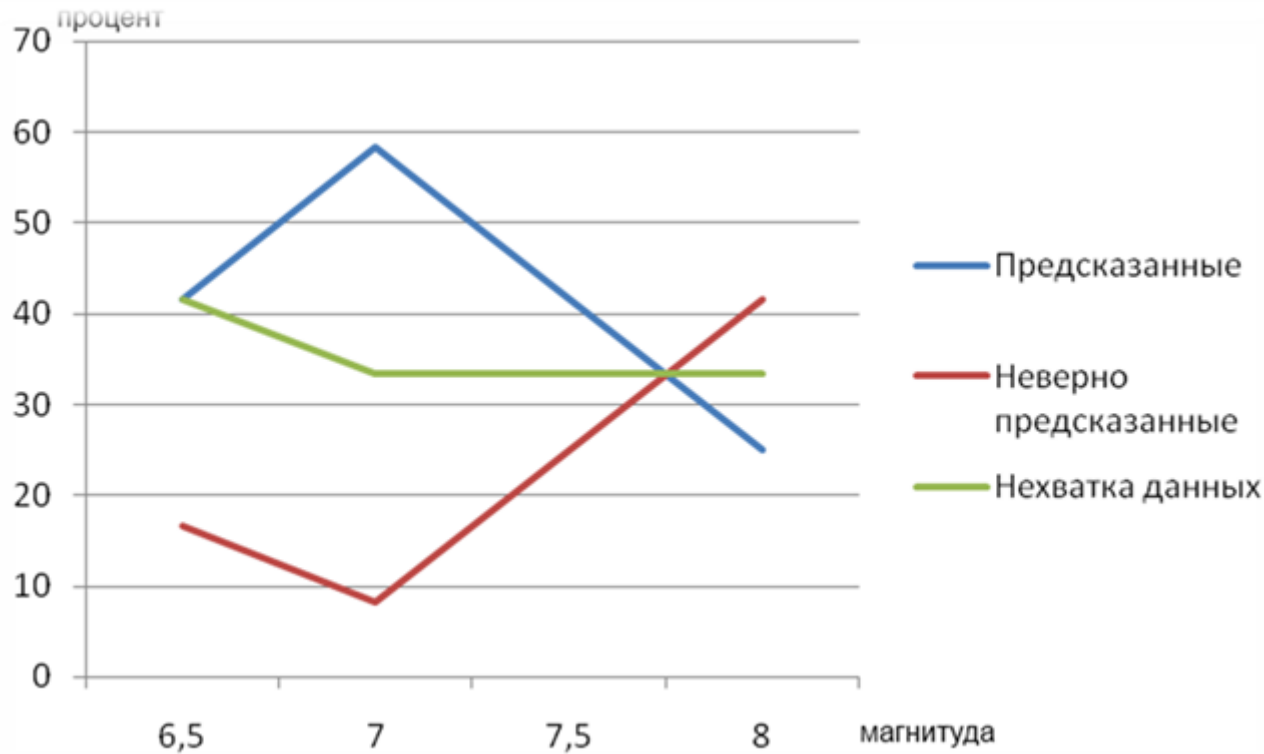
Кластеризация алгоритмом X-means землетрясений Новой Зеландии.

Магнитуда	Предсказанные		Неверные показатели		Нехватка данных	
	шт.	%	шт.	%	шт.	%
6.5	5	41.7	2	16.7	5	41.7
7	7	58.3	1	8.3	4	33.3
7.5	5	41.7	3	25	4	33.3
8	3	25	5	41.7	4	33.3
Все	20	41.7	11	22.9	17	35.4

Результаты тестирования данных Новой Зеландии



Результаты



Результаты работы алгоритма M8 на данных Новой Зеландии



Лучшие показатели для магнитуды 7



Выводы



Учебная часть

Работа имела более учебный характер, демонстрацию возможностей DM в геологии, для прикладного применения

Научная составляющая

Проведено комплексное тестирование алгоритма M8 на территории Новой Зеландии с предварительной обработкой данных

Рекомендации

Сервисы по хранению сейсмических каталогов нуждаются в внедрении алгоритмов предобработки и очистки данных





Спасибо за внимание!

