



Некоторые аспекты создания систем поддержки принятия решений на базе нейросетей

Забабахинские научные чтения
Международная конференция
29 мая – 2 июня 2023 г.
Секция №7

Балута В.И.
ИПМ им. М.В.Келдыша РАН

Предпосылки

Указом Президента РФ от 10 октября 2019 г. №490 "О развитии искусственного интеллекта в Российской Федерации" утверждена «Национальная стратегия развития искусственного интеллекта на период до 2030 года», в которой сделан акцент на необходимости интенсивного развития и внедрения средств искусственного интеллекта в систему государственного управления.

Характерные особенности ЧС

- Редкие и труднопрогнозируемые события
- Уникальность развития обстановки
- Высокая динамичность смены ситуации
- Недостаточный объем наблюдений
- Невозможность натуральных экспериментов
- Значительная неопределенность обстановки
- Большая вероятность ошибочных решений
- Высокая цена неверных решений
- ...

Сравнение подходов СППР

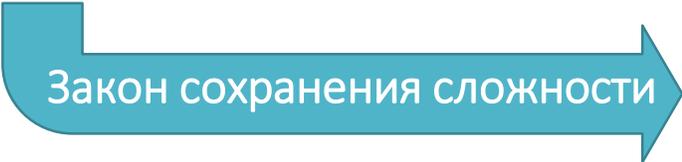
Экспертные подходы

- высокие требования к уровню квалификации специалистов
- сложности оценки обстановки
- высокая вероятность ошибочных решений

Мат.моделирование

- проблема задания корректных исходных данных в экстренных ситуациях
- большие ресурсные затраты на получение адекватных прогнозов

Закон сохранения сложности



ИИ на базе нейросетей

- возможность применения в условиях неопределенности
- уточнение прогноза в реальном времени по мере поступления новых данных
- проблема обучения нейросетей ввиду отсутствия требуемого объема реальных данных по развитию ЧС

Данные для обучения нейросетей

- Использование реальных данных наблюдений, полученных из различных источников
- Использование косвенных и сопутствующих данных
- Использование синтетических данных

Форматы:

- Изображения
- Текстовые файлы
- Числовые массивы

Подготовка натуральных данных

- Определение источников и организация получения данных
- Анализ и отбор данных, содержащих необходимую информацию
- Создание инструментария для разметки
- Разметка данных (выделение информационного содержания)
- Присвоение меток выделенным объектам

Формирование синтетических данных

Синтетическими называют данные, полученные любыми способами помимо натуральных наблюдений.

В ряду таких способов:

- аугментация (размножение) натуральных данных путем применения к ним различных преобразований;
- симуляция исследуемых процессов на базе физико-математических моделей;
- генерация случайных данных по статистическим моделям, учитывающим законы распределения реальных данных.

Способы аугментации изображений

Формальные преобразования:

1. Геометрические (аффинные, проективные и т.п.);
2. Яркостные или цветовые;
3. Преобразование или подмена фона;
4. Внесение характерных искажений (блики, шумы, размытия и т.п.);
5. Случайные комбинации различных преобразований.

Способы аугментации числовых массивов

При обработке числовых массивов применяют:

- методы устранения аномальных выбросов (ошибочных значений);
- методы заполнения пропусков в данных;
- методы размножения (увеличения объема) путем генерации сходных по каким-то критериям данных;
- методы дистилляции (целенаправленного уменьшения объема) данных;
- симуляция дополнительных данных по моделям.

Способы аугментации текста

При аугментации текстов малоприменимы формальные преобразования, так как важно сохранять смысловое содержание. «Высшим пилотажем» считается перефразирование с заменой слов и построения фраз без потери смысла.

В числе методов:

- замены (синонимами, аббревиатурами, типовыми обозначениями ...);
- перестановки (слов, оборотов, фраз, абзацев...);
- зашумление (опечатки, случайные вставки или изъятия, нарушение пунктуации...);
- перефразирование с помощью автоматических переводчиков;
- трансформация синтаксического дерева.

Сочетание технологий

Экспертные подходы:

- выделение ключевых параметров, определение источников информации;
- отчеты должностных лиц с разбором событий и оценкой мер реагирования;
- методики и рекомендации на основе обобщения наблюдавшихся событий.

Методы симуляции:

- разработка математических моделей для описания развития явлений и событий;
- проведение численных экспериментов на базе разработанных моделей.

Косвенные и дополнительные данные:

- динамика информации в социальных сетях;
- дополнительные описания, изображения, реакции свидетелей.

Порядок создания СППР на базе ИС

- Классификация кризисных и чрезвычайных ситуаций (КЧС), требующих оперативного реагирования
- Выделение характерных особенностей и информационных параметров КЧС
- Определение источников и способов получения информации по КЧС
- Сбор и консолидация натуральных данных, включая косвенные и сопутствующие
- Построение математических моделей КЧС и их валидация
- Проведение численных экспериментов и накопление результатов
- Формирование обучающей базы синтетических данных
- Первичное обучение нейросети на синтетических данных
- Дообучение нейросети на реальных данных
- Проверка применимости в игровых ситуациях