




Системы принятия решений: пара полезных свободных инструментов

Использование для извлечения неочевидных
знаний



Проблема управления знаниями в организации.

- Знания неочевидны
- Знания персонифицированы
- Знания мимолетны



Проблема: неструктурированные текстовые знания

Поток информации, по которому надо принять решение вида “да-нет”.

Примеры потока:

- Разнообразные новости
- Страницы Wiki
- Документация
- Служебные документы



Применяем `spamassassin` не по назначению.

Накапливаем исторический массив текстовых данных, сопровождая его результирующими оценками

Формируем поток обучения и направляем его в `spamassassin`, причем так, что “плохо” влияющие на параметр знания явно помечаются как спам! Полученный анализатор будет классифицировать новые данные и выдавать им оценку




Экспертные знания

Вторая задача - выделение знаний мимолетных, неочевидных и известных экспертам.

Обычно решается экспертными системами или языками декларативного программирования

Проблемы:

- Сложно! Дорого!
- Плохо привязываются к процедурам



Возможное решение проблемы - использовать Pyke - экспертную систему на питоне

- pyke.sf.net
- Позволяет определять правила в стиле декларативного программирования с бэктрекингом
- Правила хранятся отдельно и легко модифицируются экспертами
- Позволяет включать процедурный код на питоне.



Было использовано для

Исследование влияния новостей на цены акций
Обработка пакетной базы grm с целями
классификации их по степени влияния на
безопасность.