

УДК 519.254

КЛАССИФИКАТОРЫ ДИНАМИКИ КУРСА ВАЛЮТ НА ОСНОВЕ НОВОСТЕЙ И АНАЛИТИЧЕСКИХ ОБЗОРОВ РЫНКА ФОРЕКС

Т. Гараев¹, М. Александров^{1,2}, О. Кошулько³

¹Российская Академия Народного Хозяйства
и Государственной Службы при Президенте РФ, Россия

²Автономный Университет Барселоны, Испания

³Институт Кибернетики имени В. М. Глушкова, Украина

gtimurf@yandex.ru, malexandrov@mail.ru, koshulko@gmail.com

Показано можливість виконувати автоматичний прогноз котирування пари Євро/Доллар на основі текстової інформації, пов'язаної з ринком Форекс. Пропонована методика має такі три кроки: 1) відбір лінгвістичних індикаторів (ключових слів), які використовуються як змінні у прогнозних моделях; 2) ручне маркування текстів для навчання класифікаторів; 3) побудова класифікаторів. У класифікаторах використовуються різні комбінації класів: зростання, спадання, незмінність, відсутність зростання, відсутність спадання. Для побудови класифікаторів використовуються два підходи: індуктивне моделювання (пакет GMDH Shell) та регресійний аналіз (пакет Eviews). Результати експериментів показують перевагу класифікаторів, побудованих за допомогою МГУА, а також можливість використовувати запропоновану методику як корисне доповнення до існуючих чисельних методів.

Ключові слова: індуктивне моделювання, регресійний аналіз, ринок Форекс, аналіз думок

The possibility to make automatic forecasting for rating the pair Euro/Dollar on the basis of textual information related to Forex market is shown. The proposed technique includes the following three steps: 1) selecting linguistic indicators (keywords) used as variables in models; 2) manual marking of texts for training classifiers; 3) constructing classifiers. Different combinations of classes are used in the classifiers: growth, fall, invariability, non-growth, non-fall are used in the classifiers. To create the classifiers, two approaches are used: inductive modeling (GMDH Shell package) and regression analysis (Eviews package). Results of experiments show advantage of the classifiers built by GMDH and also possibility to use the proposed technique as a useful addition to existing numerical methods.

Keywords: inductive modeling, regression analysis, Forex market, opinion analysis

Показана возможность выполнять автоматический прогноз котировки пары Евро/Доллар на основе текстовой информации, связанной с рынком Форекс. Предлагаемая методика включает в себя следующие три шага: 1) отбор лингвистических индикаторов (ключевых слов), которые используются как переменные в прогнозных моделях; 2) ручную маркировку текстов для обучения классификаторов; 3) построение классификаторов. В классификаторах используются разные комбинации классов: рост, падение, неизменность, отсутствие роста, отсутствие падения. Для построения классификаторов используются два подхода: индуктивное моделирование (пакет GMDH Shell) и регрессионный анализ (пакет Eviews). Результаты экспериментов показывают преимущество классификаторов, построенных с помощью МГУА, а также возможность использовать предложенную методику как полезное дополнение к существующим численным методам.

Ключевые слова: индуктивное моделирование, регрессионный анализ, рынок Форекс, анализ мнений

Введение

Игроки рынка Форекс используют различные стандартные математические и нестандартные эвристические модели для прогноза. На страничках Интернета можно легко найти множество методик построения таких моделей. При этом даются общие рекомендации, но «дьявол сидит в деталях». Как правило, игроки не публикуют эти детали, держа их в секрете. И это вполне оправдано, поскольку, модели, известные множеству игроков, сразу перестают быть выигрышными.

В такой ситуации было бы полезно применять методику, которая бы не зависела от знаний и опыта конкретного игрока, а использовала бы интегральный опыт тех, кто имеет дело с Форекс. Этот опыт отражается в новостях и аналитических обзорах рынка Форекс. В статье на конкретном примере прогноза поведения валютной пары Евро/Доллар мы показываем возможность использования этой текстовой информации.

В работе мы строим 4 типа классификаторов. Соответствующие классы и их обозначения представлены в таблице 1. Сравниваются 2 конкурирующих подхода: индуктивное моделирование и традиционный регрессионный анализ. Индуктивное моделирование (ИМ) реализуется с помощью пакета GMDH Shell [1]. Регрессионный анализ (РА) выполняется средствами статистического пакета EViews [2]. Эксперименты проводятся на реальных текстах, полученных с одного из официальных сайтов, где имеется архив упорядоченных по времени новостей и обзоров.

Табл.1.

Типы классификаторов

<i>Тип</i>	<i>Классы</i>	<i>Обозначения классов</i>
1	Рост, Падение	{+1, -1}
2	Рост, Неизменность, Падение	{+1, 0, -1}
3	Рост, Отсутствие роста	{+1, 0}
4	Падение, Отсутствие падения	{-1, 0}

Оставшиеся разделы статьи организованы следующим образом. В разделе 1 описываются лингвистические ресурсы – корпус и лексика – для построения классификаторов. В разделе 2 мы строим классификаторы и проводим эксперименты с использованием МГУА и РА. Заключение содержит выводы по проведенному исследованию. Статья представляет собой расширенный вариант ранее опубликованной работы авторов [3].

2. Лингвистические ресурсы

2.1 Корпус документов

Новости и обзоры Форекс были взяты с сайта <http://www.dailyfx.com/>. Данный сайт своевременно предоставляет достаточно полную информацию о

новостях/обзорах рынка Форекс и пользуется доверием игроков. В приложении 1 представлены два примера новостей

В качестве текстовой единицы берутся тексты, появившиеся в течение 2-х дней до момента торгов. Обычно, одна текстовая единица содержит 3-5 новостей/обзоров. Основные характеристики корпуса документов представлены в таблице 2.

Табл.2.

Характеристики корпуса документов

	<i>Значение</i>
Количество документов	89
Средняя длина текстовой единицы (слов)	513

Таблица 3 иллюстрирует пример исходных данных. Здесь в первом и втором столбцах показана дата торгов и имя текстовой единицы, в которой собраны новости и обзоры за 2 дня, предшествующие торгам. В предпоследнем и последнем столбцах показаны результаты котировок и классы, которым принадлежит динамика изменений. Следует прояснить принадлежность классу 'неизменность'. Очевидно, что вероятность получить точное совпадение котировок равна нулю. Мы условно принимаем, что имеет место постоянство котировок, если изменение котировок меньше некоторого порога. Значение порога выбирается экспертом.

Табл.3.

Исходные данные для обучения модели

<i>Дата</i>	<i>Документы</i>	<i>Котировка Евро/Доллар</i>	<i>Изменение</i>
13.03.2012	Text-1	1,25	
15.03.2012	Text-2	1,15	Падение
17.03.2012	Text-3	1,24	Рост
19.03.2012	Text-4	1,22	Неизменность
21.03.2012	Text-5	1,26	Рост

Структура корпуса документов для разных классов представлена в таблице 4. Следует заметить, что для бинарного классификатора {Рост, Падение} тексты, относящиеся к классу Неизменность были распределены между классами Рост и Падение.

Табл.4.

Структура корпуса документов

<i>Тип</i>	<i>Рост</i>	<i>Неизменность</i>	<i>Падение</i>	<i>Отсутствие роста</i>	<i>Отсутствие падения</i>	<i>Всего</i>
Тип 1	43	–	46	–	–	89
Тип 2	30	24	35	–	–	89
Тип 3	30	–	–	59	–	89
Тип 4	–	–	35	–	54	89

2.2 Лингвистические переменные и параметризация текстов

Лингвистическими переменными являются ключевые слова и выражения (мы будем их также называть ключевыми словами), которые позволяют трансформировать текстовые документы в их векторную форму. Алгоритм построения списка ключевых слов состоит из двух шагов

- 1) Автоматический отбор слов программой на основе некоторого критерия селекции;
- 2) Ручная коррекция полученного списка слов. Эксперт корректирует список, удаляя лишние слова и добавляя новые

Для отбора слов мы используем программу LexisTerm, которая имеется в свободном доступе [4]. В этой программе применяется так называемый критерий специфичности. Здесь понятие специфичность определяет отношение частоты встречаемости термина в данном корпусе к частоте встречаемости этого же термина в некотором эталонном корпусе. Специфичность задается числом $K \geq 1$. В качестве эталонного корпуса используется Британский Национальный корпус. Его лексика, называемая Генеральной лексикой английского языка, имеется в свободном доступе. Таким образом, входными данными для указанной программы являются: исходный корпус текстов, Генеральная лексика английского языка, и пороговое значение K для критерия специфичности. В приложении 2 показан интерфейс программы LexisTerm.

Мы сравнивали списки, получаемые программой при $K=2, 5, 10, 20, 50$. Оказалось, что при значениях $K < 10$, получаемые списки слов включали слишком много незначимых терминов. При $K > 10$ наоборот – терялось много значимых терминов. Именно поэтому в работе мы использовали значение порога $K=10$. «Логарифмический» характер значений K в эксперименте со списками слов был выбран по причине того, что лишь при таких больших шагах в значениях порога можно было наблюдать существенные изменения в количестве отобранных терминов. Этот феномен был отмечен в работе [4].

Для значения порога $K=10$ были отобраны 155 ключевых слов. Наиболее частыми ключевыми словами в списке оказались: *inflation, german, figure, economic, account, consume, trade, changes, goods, growth, rates, prices*, и т.д. Эти слова можно увидеть в правом окне программы, на скриншоте. Затем этот список был скорректирован экспертом (одним из авторов статьи), и итоговый список содержал уже 96 слов. Эти ключевые слова можно теперь назвать лингвистическими переменными.

Для параметризации мы использовали программу DocTerm. Программа имеется в свободном доступе, ее описание будет представлено в [5]. Входными данными для программы являются исходный корпус документов и список терминов. На выходе программы мы имеем матрицу документ-термин. Эта матрица отражает частоту вхождения терминов списка в документы корпуса. Программа позволяет работать с терминами, включающими 1,2 и 3 слова. Она может быть использована с текстами, написанными в латинице и кириллице. В

ней также могут быть использованы коэффициенты значимости терминов, позволяющими учесть их важность для данного корпуса.

Следует сказать, что построение матрицы документ-термин для случая однословных терминов и текстов на латинице может быть легко запрограммировано на языках C++, R, и ряде других, и тогда нет необходимости использовать какие-либо другие внешние программы.

В результате параметризации была получена матрица размером 89x96, которая отражала векторную форму представления документов в пространстве терминов. Здесь 89 и 96 это число документов и терминов соответственно. Все значения частот были нормализованы к единичной длине вектора частот для каждого документа. В приложении 3 показана часть таблицы параметризованных документов.

2. Построение прогнозных моделей

2.1 Индуктивное моделирование

Для реализации индуктивного моделирования мы использовали программную оболочку GMDH Shell (GS). Это широко известный инструмент прогноза временных рядов, аппроксимации функций и классификации объектов, включающий расширенные возможности для визуализации результатов [1]. GS использует технику МГУА – метода группового учета аргументов [6,7]. В настоящее время GS включает модификации двух алгоритмов:

- Комбинаторный алгоритм МГУА
- Нейроподобный алгоритм МГУА

Для решения нашей задачи мы применяли режим классификации. В этом режиме GS использует метод One-vs-All [8]. Метод сводит многоклассовую классификацию к бинарной классификации, и, таким образом, для заданной обучающей выборки строится столько бинарных классификаторов, сколько классов мы имеем. Каждый из бинарных классификаторов определяет степень доверия к заданному классу (через точность, вероятность, и т.п.). Выбирается класс, обеспечивающий наибольшую степень уверенности. Очевидно, что если у нас имеется только два класса, то достаточно иметь один бинарный классификатор.

Исходный корпус классифицированных (маркированных) документов содержит 89 документов. Для построения модели, то есть для обучения и контроля (learning = training+control), мы взяли 80% документов всего корпуса. Это число составляет 71 документ. Оставшиеся 18 документов были использованы для проверки модели (verification). Очевидно, что такая выборка невелика, и потому полученные на ней оценки надо рассматривать, как весьма приближенные. Действительно, здесь одно ошибочное отнесение документа к чужому классу сразу изменяет точность на 6% (~1/16)

Были построены классификаторы всех 4-х типов. Тип классификатора и используемые им классы отражены в таблице 1. Точность классификации для каждого типа классификаторов представлена в таблице 5. Здесь во втором столбце указано значение точности классификатора, построенного на обучающей выборке, на данных контрольной выборки. В третьем столбце указана точность на проверочной выборке.

Для построения всех типов классификаторов был использован комбинаторный МГУА

Табл.5.

Точность классификаторов (ИМ)

<i>Тип</i>	<i>Обучение- Контроль</i>	<i>Проверка</i>
Тип 1	78,9%	61,1%
Тип 2	67,6%	44,4%
Тип 3	71,8%	72,2%
Тип 4	67,6%	61,1%

Ниже представлены уравнения классификаторов. В этих уравнениях переменная Y в левой части есть значение классификатора, которое сравнивается с порогом при принятии решения. Переменные w_i в правой части являются частотами соответствующих лингвистических индикаторов в рассматриваемом документе.

Тип 1 {Рост = 1, Падение = -1}

$$Y = 0.660 - 10.204w_{16} - 75.236w_{29}w_{55} + 619.791w_{37}w_{45}$$

Правило принятия решения: $Y \geq 0$ выбирается класс Рост, иначе выбирается класс Падение

Тип 2 {Рост = 1, Неизменность=0, Падение = -1}

$$Y_{+1} = 0.247 + 7.034w_8w_{33} - 38.189w_{20}w_{33} + 7.232w_{52}$$

$$Y_0 = 0.043 + 23.794w_9w_{80} - 6.498w_{11}w_{73} + 71.898w_{19}^2$$

$$Y_{-1} = 0.211 + 90.171w_{16}w_{20} + 81.230w_{29}w_{55} + 37.955w_{53}w_{54}$$

Правило принятия решения: выбирается класс, имеющий максимальное значение Y

Тип 3 {Рост = 1, Отсутствие роста = 0}

$$Y = 0.247 + 7.034w_8w_{33} - 38.189w_{20}w_{33} + 7.232w_{52}$$

Правило принятия решения: $Y \geq 0.5$ выбирается класс Рост, иначе выбирается класс Отсутствие роста

Тип 4 {Падение = -1, Отсутствие падения = 0}

$$Y = 0.789 - 90.171w_{16}w_{20} - 81.230w_{29}w_{55} - 37.955w_{53}w_{54}$$

Правило принятия решения: $Y \geq -0.5$ выбирается класс Отсутствие падения, иначе выбирается класс Падение

В качестве примера работы программы в приложении 4 представлен скриншот GS для классификатора {Рост, Отсутствие роста}. На скриншоте в верхней таблице детально представлены результаты работы программы на контрольной выборке, а в нижней таблице – на проверочной.

2.2 Регрессионный анализ

Для проведения регрессионного анализа мы использовали известный статистический пакет EViews (Econometric Views). Пакет позволяет строить линейные регрессионные модели, используя метод наименьших квадратов и относительно большой набор статистических критериев [2].

Способ построения классификатора на основе регрессионного уравнения известен. Тем не менее, мы вновь приведем его описание, чтобы избежать пояснений некоторых величин, используемых далее в тексте статьи.

- 1) Построение уравнения регрессии. Здесь классам присваиваются последовательные числовые обозначения в заданном интервале значений, например, в интервале $[-1,1]$. Тогда при бинарной классификации на {Рост, Падение} классы получают значения $\{1, -1\}$, а при классификации на 3 класса {Рост, Неизменность, Падение} классы получают значения $\{1, 0, -1\}$ соответственно. Указанные значения используются в левой части уравнения регрессии при проведении РА
- 2) Установка порогов. Число порогов для уравнения регрессии на единицу меньше числа классов и, таким образом, при бинарной классификации устанавливается один порог. Тогда при превышении этого порога объекту будет присоедена принадлежность 1-му классу, в противном случае – 2-му классу. При классификации на три класса необходимо выбрать два порога. Тогда, если значение уравнения будет больше большего порога или меньше меньшего порога, то объект будет отнесен, соответственно, к 1 или 3 классу. В противном случае, он будет отнесен ко второму классу.

Примечание. Выбор порогов выполняется визульно на основе наблюдения распределения значений уравнения регрессии для каждого класса. Этот процесс можно автоматизировать, поручив программе перебор вариантов.

При проведении РА мы встретили естественную трудность, которая состояла в том, что количество наблюдений (89 документов) меньше числа переменных (96 терминов). Очевидно, что в таких условиях РА невозможен. Чтобы обойти эту трудность мы использовали только те переменные, которые были отобраны в процессе индуктивного моделирования в оболочке GS. Вот эти переменные: $\{w_8, w_9, w_{11}, w_{16}, w_{20}, w_{29}, w_{33}, w_{52}, w_{53}, w_{54}, w_{55}\}$. Тогда уравнение регрессии, которое требуется построить, принимает вид:

$$Y = a_8w_8 + a_9w_9 + a_{11}w_{11} + a_{16}w_{16} + a_{20}w_{20} + a_{29}w_{29} + a_{33}w_{33} + a_{52}w_{52} + a_{53}w_{53} + a_{54}w_{54} + a_{55}w_{55}$$

Здесь, Y в левой части уравнений есть значение уравнения регрессии, переменные w_i в правой части уравнений являются частотами соответствующих лингвистических индикаторов в рассматриваемом документе, a_i искомые коэффициенты

Как и ранее при ИМ сейчас в РА для построения модели мы использовали 71 документ (learning=training), а для ее проверки 18 документов (verification).

При помощи EViews мы строили модели для всех 4-х типов классификаторов. Однако, состоятельные модели, то есть модели со значимыми коэффициентами уравнения регрессии удалось получить только для комбинированных классификаторов: {Рост, Отсутствие роста}, и {Падение, Отсутствие падения}. После построения уравнения были рассчитаны оптимальные значения порогов для каждого классификатора. Расчет проведен в программе на VBA путем простого перебора. Так, для классификатора {Рост, Отсутствие роста} значение порога оказалось равным $U=0,387$. Для классификатора {Падение, Отсутствие падения} это значение оказалось равным $U=-0,344$. В приложении 5 показано распределение значений уравнений регрессии для классификатора {Рост, Отсутствие роста} с обозначениями классов {1,0}. Легко видеть, что указанный порог $U=0,387$ вполне подходит для различения указанных классов. Все результаты сведены в таблицу 6.

Табл.6.

Точность классификаторов (РА)

<i>Тип</i>	<i>Обучение</i>	<i>Проверка</i>
Тип 1	–	–
Тип 2	–	–
Тип 3	72,9%	68,4%
Тип 4	59,0%	78,9%

Ниже представлены уравнения классификаторов. В этих уравнениях переменная Y в левой части есть значение уравнение регрессии, которое сравнивается с порогом при принятии решения. Переменные w_i в правой части являются частотами соответствующих лингвистических индикаторов в рассматриваемом документе.

Тип 3 {Рост=1, Отсутствие роста=0}

$$Y = 0.205 + 1.106w_8 - 4.614w_{20} + 4.828w_{52}$$

Правило принятия решения: $Y \geq 0.387$ выбирается класс Рост, иначе выбирается класс Отсутствие роста

Тип 4 {Падение=-1, Отсутствие падения=0}

$$Y = -4,578w_{20} - 3,677w_{54} - 3,677w_{55}$$

Правило принятия решения: $Y \geq -0.344$ выбирается класс Отсутствие падения, иначе выбирается класс Падение

Заключение

Результаты работы сводятся к следующему:

- Предложена новая методика предсказания поведения валютной пары, основанная на анализе новостей и аналитических обзоров рынка Форекс;
- Проведено сравнение двух подходов к построению классификаторов текстов, при этом техника ИМ показала в целом лучшие результаты, чем РА. Здесь следует иметь ввиду: возможность построения всех типов классификаторов в ИМ, использование результатов МГУА в РА, и крайнюю ограниченность проверочной выборки
- Предложенная методика не может заменить существующие стратегии игроков на рынке Форекс, но может быть полезным дополнением к ним.

В будущем мы планируем:

- оценить зависимость качества классификации от состава и размера списка лингвистических индикаторов;
- применить методику к другим валютным парам.

Литература

1. Ресурс GMDH Shell, <http://gmdhshell.com>
2. Ресурс EViews, <http://www.eviews.com>
3. Garaev, T., Alexandrov, M., Koshulko, O.: Text classifier as a tool for short-term forecast of currency rates. Proc. of 4th Intern. Conf. on Inductive Modeling (ICIM-2013), NAS of Ukraine, Prague Technical University, 2013, pp. 261-266
4. Lopez, R., Alexandrov, M., Barreda, D., Tejada, J.: Lexistern – the program for term selection by the criterion of specificity. In: Artificial Intelligence Applications to Business and Engineering Domain. Rzeszov-Sofia, ITNEA Publ., No.24, 2011, pp.8-15
5. Александров, М., и др: Инструменты анализа данных для Интернет исследований. Препринт РАНХиГС, 2014 (готовится к печати, апрель 2014)
6. Ивахненко А.: Индуктивный метод самоорганизации моделей сложных систем. Киев, Наук. думка, 1982. 296 с.
7. Ивахненко А., Степашко, В.: Помехоустойчивость моделирования. Киев, Наукова думка, 1985, 216 с.
8. Википедия, http://en.wikipedia.org/wiki/Multiclass_classification

Приложения

П1. Примеры документов

№	Документ
1.	Measures changes in sales of the German retail sector. Given that consumption makes up a significant portion of German GDP, the Retail Sales figure can act as an indicator of domestic demand. High or rising Retail Sales may spur German consumption, translating into economic growth. However, uncontrolled growth runs the risk of inflationary pressures. Since Germany is a large part of the Euro-zone, German figures may have some impact on the market. The headline figure is expressed in percentage change in the value of sales.
2.	A monthly estimate of inflation in the Australian economy. The report replicates the methodology used by the Australian Bureau of Statistics to calculate quarterly CPI, striving to correspond closely with official government numbers. The figure is important because it acts as a timelier indicator of inflation, coming out monthly instead of the quarterly CPI figures. Released one day before interest rate decisions are made, the figure may influence RBA considerations for rate hikes or reductions. As with any gauge of inflation in Australia, a high value in the figure is bullish for the Australian dollar since real inflationary pressures are usually met by the Reserve Bank of Australia with bullish rate hikes. The headline figure is the month'over'month or annualized inflation rate.

П.2 LexisTerm (скриншот)

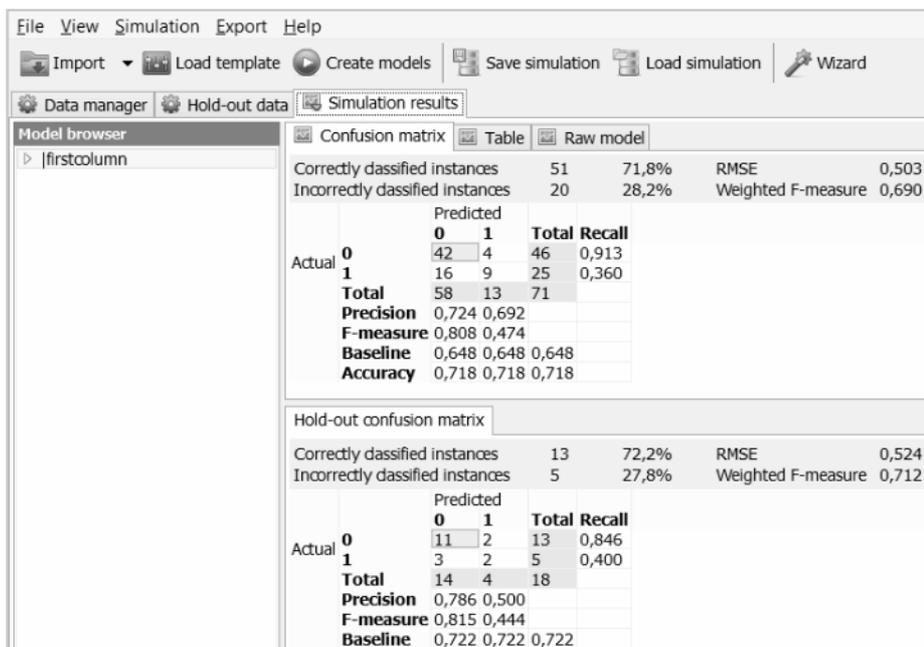
The screenshot displays the LexisTerm software interface, which is divided into three main sections: CORPUS, GENERAL LEXICON, and RESULTS. Each section contains a frequency list of words and their corresponding counts. The CORPUS section shows the most frequent words in the document, such as 'the' (85,588,799) and 'of' (40,977,393). The GENERAL LEXICON section shows the frequency of words in the general lexicon, with 'the' (79,179,369) and 'of' (41,297,969) being the most frequent. The RESULTS section shows the frequency of words in the results, with 'economic' (0,006666) and 'account' (0,006198) being the most frequent. The interface also includes fields for Folder Path, File Path, Output File Path, Normalization options, and a Run button.

CORPUS		GENERAL LEXICON		RESULTS		
Word	Frequency	Word	Frequency	Word	Frequency	Documents
the	85,588,799	the	79,179,369	economic	0,006666	78
of	40,977,393	of	41,297,969	account	0,006198	38
in	31,390,734	and	33,240,906	consumer	0,006064	67
and	25,549,561	a	25,131,609	trade	0,005953	31
a	24,992,197	in	24,250,766	economy	0,00593	76
is	20,421,813	to	19,345,481	current	0,005886	39
to	18,036,296	is	13,934,454	changes	0,005841	74
as	11,860,704	to	11,394,288	services	0,005618	61
for	11,548,58	for	11,093,974	goods	0,004905	66
eurozone	10,099,434	it	9,527,667	confidence	0,00457	27
that	9,720,426	be	8,491,876	index	0,004504	80
are	9,408,302	was	8,308,766	growth	0,004481	70
inflation	7,268,025	on	7,847,778	percentage	0,004281	77
on	7,045,08	with	7,801,738	rates	0,004281	57
german	6,889,018	that	7,494,295	sector	0,004236	42
figure	6,777,545	by	7,405,72	prices	0,004191	56
economic	6,666,072	are	6,621,632	germany	0,004035	65
account	6,197,886	's	5,868,51	higher	0,00379	67
Leum	6,175,592	this	5,666,82	value	0,003589	50

П.3 Параметризованный корпус текстов (часть данных)

	A	B	C	D	E	F	G	H
	DOCs	Growth Constancy Fall	Category	W 1	W 2	W 3	W 4	W 5
1								
2	Text 1	-1	C	0,00	0,00	0,00	0,00	
3	Text 2	-1	C	0,03	0,00	0,00	0,00	
4	Text 3	-1	C	0,16	0,00	0,00	0,00	
5	Text 4	-1	C	0,49	0,02	0,07	0,26	
6	Text 5	-1	C	0,00	0,00	0,00	0,00	
7	Text 6	-1	C	0,05	0,00	0,00	0,00	
8	Text 7	-1	C	0,00	0,05	0,03	0,21	
9	Text 8	1	A	0,00	0,00	0,00	0,00	
10	Text 9	-1	C	0,00	0,00	0,00	0,00	
11	Text 10	-1	C	0,36	0,04	0,04	0,11	
12	Text 11	1	A	0,00	0,00	0,00	0,00	
13	Text 12	-1	C	0,00	0,00	0,00	0,00	
14	Text 13	1	A	0,08	0,00	0,00	0,00	

П.4 GMDH Shell, классификатор {Рост, Отсутствие роста}



П.5 Значения регрессии на классах {Рост, Отсутствие роста}

