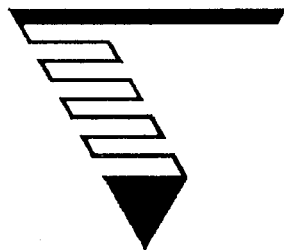


РОССИЙСКАЯ АКАДЕМИЯ НАУК  
Вологодский научно-координационный центр



Препринт

*А.Е.Прокофьев*

**ФАКТОРНЫЙ АНАЛИЗ В РАМКАХ МОДЕЛИРОВАНИЯ  
ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ**

Вологда  
1997

## АННОТАЦИЯ

Факторный анализ- один из разделов многомерного статистического анализа, объединяющий вычислительные методы , которые в ряде случаев позволяют получить компактное описание исследуемых явлений на основе обработки больших массивов информации. Отличительной особенностью факторного анализа является то, что он не опирается на заранее заданный “априорный” перечень факторов, влияющих на изучаемые переменные, а наоборот, при соблюдении определенных правил и предосторожностей помогает обнаружить наиболее важные из этих факторов, причем скрытые. В настоящей работе делается попытка решить одну из задач факторного анализа с помощью моделирования искусственной нейронной сети. Нейронная сеть “обучается” на основе реальных статистических данных, отражающих динамику основных макроэкономических показателей для Вологодской области за 1995-1996 гг, и с помощью определения чувствительности входов нейросети проводится анализ факторов, влияющих на поступление налогов в бюджетную систему.

# ФАКТОРНЫЙ АНАЛИЗ В РАМКАХ МОДЕЛИРОВАНИЯ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ

(на примере макроэкономической статистики для Вологодской области)

к.ф.-м.н. Прокофьев А.Е., Вологодский научно-координационный центр РАН

В работе [ 1 ] описана нейросетевая модель для краткосрочного прогнозирования основных показателей социально-экономического развития региона. Модельные расчеты основаны на использовании системы макроэкономических показателей за ряд месяцев 1995-1996 гг., причем показатели за каждую пару смежных месяцев рассматриваются в качестве обучающего примера (рис.1). Программными средствами устанавливаются такие взаимосвязи между входами и выходами, которые в наибольшей степени соответствуют реальному положению вещей

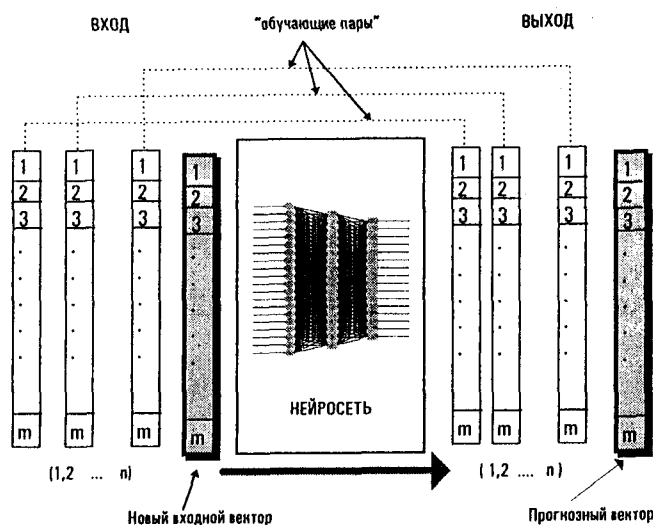


Рис. 1 Схема модельных расчетов при решении задачи прогнозирования

(аналог работы нейронов в режиме обучения). Найденные структуры переходных функций используются для получения краткосрочных прогнозов (на 1 месяц вперед), основываясь на данных конца базового периода, затем - на результатах найденного прогноза и т.д. В работе [ 2 ] приведены результаты сопоставления прогнозных оценок (прогноз был сделан в августе 1996 г.) и фактических значений основных макроэкономических показателей для Вологодской области в сентябре-декабре 1996 г. Данные [ 2 ] показывают, что для большинства макроэкономических показателей прогнозные оценки незначительно отличаются от фактических, что свидетельствует о высокой степени достоверности прогнозов, получаемых в рамках нейросетевого моделирования.

Вместе с тем остался неисследованным один вопрос, имеющий в какой-то степени дискуссионный характер. Суть этого вопроса сводится к следующему. Несмотря на то, что правильно обученная нейронная сеть хорошо решает задачи прогнозирования, ее внутреннее наполнение (матрицы весовых коэффициентов), фактически не имеет содержательной интерпретации. В то же время при построении моделей для экономического прогнозирования (в частности эконометрических моделей) именно изучение структуры взаимосвязей между отдельными переменными, интерпретация получаемых коэффициентов моделей как раз и являются основными моментами работы. Нейронная сеть остается для исследователя-экономиста "черным ящиком", и невозможно объяснить почему она классифицировала входной образ так, а не иначе.

Искусственные нейронные сети строятся по примитивному подобию мозга человека, поэтому определенные аналогии с ним уместны [3]. Попробуйте объяснить, почему двухлетний ребенок достаточно уверенно распознает на

фотографиях свою бабушку в любой одежде и с любым выражением лица. Заметим, что бабушка живет в другом городе, и ребенок никогда не общался с ней “в живую”. Понятно, что перед этим в течение некоторого времени проходил процесс обучения: изображению бабушки на разных фотографиях приводился в соответствие набор звуков, на слух воспринимаемый посторонним наблюдателем как “ба-ба”. Для этой процедуры задействуется некоторая область мозга ребенка, содержащая сравнительно небольшое число нейронов, по сравнению с их общим количеством, и фактически настраиваются силы связи (весовые коэффициенты) между нейронами в так называемых синапс-зонах. Если процесс обучения прошел успешно, а это всегда можно проверить с помощью тестовых проверок, можно перейти к более смелым экспериментам с ребенком, например, показать ранее не предъявлявшуюся ему цветную фотографию бабушки с подругами (обучение проходило на черно-белых фотографиях). Окажется, что ребенок и в этом случае безошибочно остановит свой пальчик на фотографии в нужном месте и скажет “ба-ба”.

Говоря формальным языком, после прямого распространения входного сигнала (вектора), содержащего массу признаков (группа признаков одежды, группа признаков лица, группа признаков цвета и т.д.), через обученную нейронную сеть на выходе возникает сигнал, свидетельствующий о возможности или невозможности отнесения этого вектора к определенному классу. При этом, когда скоро нейронная сеть безошибочно решает задачу распознавания, то значит она сама определяет (через свои весовые коэффициенты), какие признаки являются существенными, а какие признаки оказывают слабое воздействие на выходной сигнал. В нашем примере с ребенком очевидно, что признаки одежды и признаки цвета являются малосущественными, а признаки лица - существенными. Из этого также следует еще один важный вывод: поскольку при обучении прямо не указывалось, какие признаки являются существенными, но процесс обучения состоялся, и задача распознавания решается верно, то следовательно информация о существенности каждого из признаков косвенным образом содержится в самих обучающих примерах.

В случае прогнозирования показателей экономического развития рассматриваемая задача может быть сформулирована так. Имеется набор входных векторов, компонентами которых являются макроэкономические показатели и набор соответствующих этим векторам скалярных величин. Можно ли в рамках моделирования нейронной сети определить, какие макроэкономические показатели оказывают наиболее сильное воздействие на выходную величину, и если можно, то имеет ли получаемый результат “экономический смысл”.

Таким образом, в данной работе делается попытка подойти к проблеме изучения структуры взаимосвязей между отдельными переменными, имеющей важное значение при построении эконометрических моделей, несколько с другой стороны. Не следует пытаться понять структуру этих взаимосвязей на основе просто анализа матриц весовых коэффициентов. Проще установить эту взаимосвязь на основе анализа прямого распространения входных сигналов через обученную нейронную сеть в целом, давая им для этого последовательно некоторые приращения. Практическая значимость решения данной задачи заключается в том, что появляется возможность оценивать смену тенденций, происходящую от того или иного сдвига в структуре макропоказателей. В конечном счете это позволит более глубоко оценивать позитивный или негативный характер такого рода сдвигов, поскольку неустойчивость и социальная несбалансированность переходного процесса остро ставят во главу угла вопросы корректировки экономической политики.

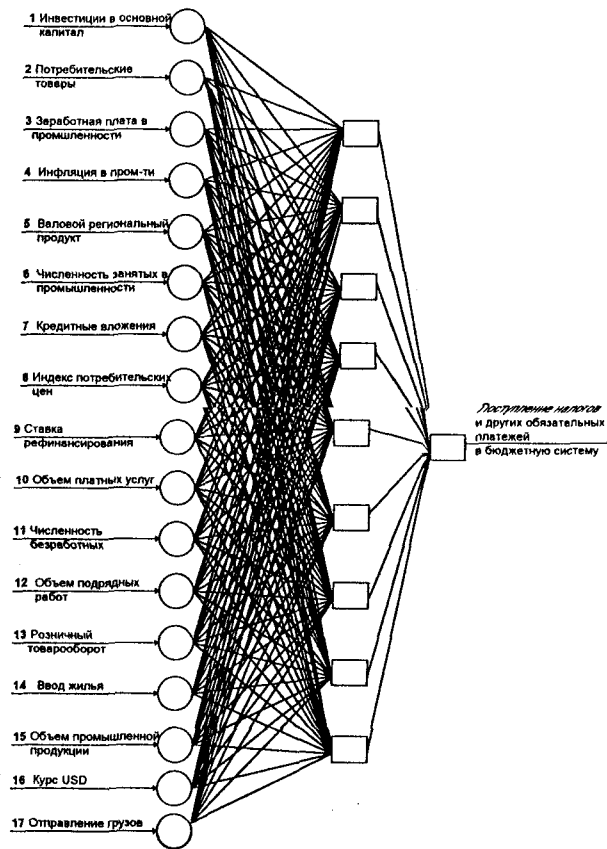


Рис. 2 Нейронная сеть для моделирования поступления налогов в бюджетную систему

бюджетом субъекта федерации определяется налоговым законодательством и законом о федеральном бюджете на текущий год; объем налогов данного вида определяется перемножением соответствующей ставки на налогооблагаемый показатель. Поэтому в точном смысле поступление налогов не нуждается в моделировании. Однако в данном случае речь может идти как об анализе факторов, влияющих на собираемость налогов, так и о выяснении латентных обобщенных факторов, которые объясняют изменения данного показателя.

На рис. 2 условно показана схема нейронной сети для моделирования по данному показателю. На первом этапе весовые коэффициенты нейронной сети прямого распространения настраиваются по алгоритму, описанному в [4], таким образом, чтобы при последовательной подаче на ее вход векторов, компонентами которых являются основные макроэкономические показатели, выходные значения были максимально близки к исследуемой величине. Отметим, что обучающие пары составлены без сдвигки по месяцам, т.е. на входе вектор макропоказателей за январь, на выходе величина налогов также за январь и т.д.

Мерой качества подгонки является коэффициент детерминации  $R^2$ :

$$R^2 = \frac{[N \sum (OP) - (\sum O)(\sum P)]^2}{[N \sum (O^2) - (\sum O)^2][N \sum (P^2) - (\sum P)^2]}, \quad (1)$$

где  $P_i$ ,  $O_i$  - текущие фактические (*Pattern*) и расчетные (*Output*) значения соответственно,  $N$  - число обучающих пар. Суммирование по всем  $i$  от 1 до  $N$ .

В качестве набора статистических данных для решения этой задачи, мы выбрали данные ежемесячной статистики для Вологодской области за 1995-1996 гг, отражающие основные макроэкономические показатели. Мы умышленно привлекли данные по области, а не по стране, поскольку нам представляется, что они более подходят для решения вышеописанной проблемы именно в силу их "неусредненности" по всем регионам страны.

Выходным параметром нейросетевой модели, на котором проводилась отработка методики, был назначен показатель, характеризующий поступление налогов и других обязательных платежей в бюджетную систему. С точки зрения регионального управленца он представляется важным. Вообще говоря, поступление налогов в бюджетную систему и их разделение между федеральным бюджетом и

На рис. 3а показана динамика поступления налогов и других обязательных платежей в бюджетную систему для Вологодской области за 1995-1996 гг. Сплошной линией показана фактическая динамика, пунктирной линией соединены расчетные значения, полученные с помощью нейросети ( $R^2=0.994$ ). Для сравнения на рис. 3б приведены аналогичные зависимости, но полученные в рамках множественной линейной регрессии ( $R^2=0.876$ ). Как видно нейронная сеть обеспечивает лучшее качество подгонки в силу нелинейности своей передаточной функции.

Интересно отметить одно обстоятельство. Посмотрев на динамику поступления налогов можно заметить, что в декабре 1996 г. собираемость налогов по сравнению с предыдущим месяцем резко повысилась (с 275 до 454 млрд. рублей). Известным фактом является то, что в декабре заседания ВЧК проходили практически еженедельно, вопрос находился на контроле у губернатора - это и обеспечило существенное улучшение собираемости налогов. Принимая во внимание, что резкий выброс может затруднить обучение нейросети, было принято решение исключить из обучающих примеров обучающую пару, соответствующую декабрю 1996 г. Тем не менее, после обучения нейронной сети, ради интереса, на ее вход был подан вектор с макропоказателями, соответствующими декабрю 1996 г. и, таким образом, получено прогнозное значение поступления налогов и других обязательных платежей в бюджетную систему в декабре 1996 г. Оно оказалось даже немного больше фактического - 467 млрд. рублей! Но откуда нейронная сеть "узнала" о заседаниях ВЧК? \*

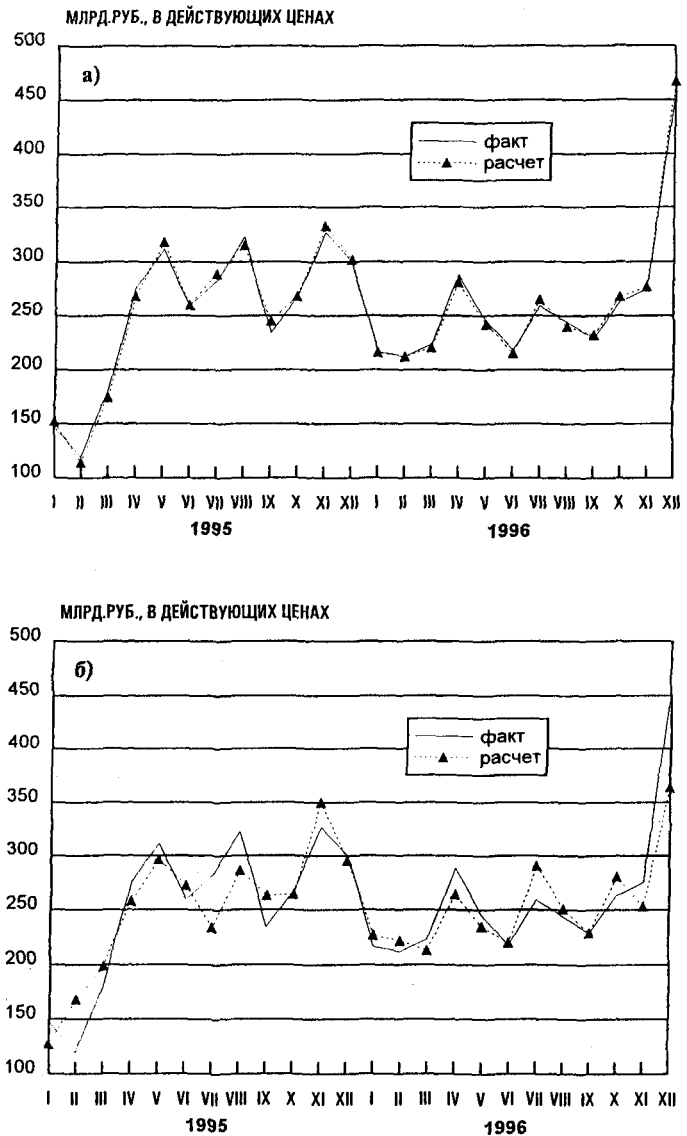


Рис.3 Динамика поступления налогов и других обязательных платежей в бюджетную систему  
а) - нейросетевое моделирование; б) - множественная линейная регрессия

\* Справедливости ради, следует указать, что данный расчет не является прогнозом в точном смысле слова, поскольку на вход подавался вектор с декабрьскими макропоказателями, на выходе поступление налогов также в декабре. Прогноз - это когда на входе вектор за декабрь, на выходе - вектор за январь. Так что, о заседаниях ВЧК нейросеть "узнала" из структуры макропоказателей за декабрь, поскольку результаты деятельности ВЧК отразились и на них.

На втором этапе расчетов проводятся следующие вычисления. На вход обученной нейросети последовательно (месяц за месяцем) подаются вектора, компонентами которых являются макроэкономические показатели данного месяца, при этом один из компонентов вектора имеет 50% -ое приращение по сравнению с фактическим значением, а все остальные компоненты вектора остаются без изменения. Для каждого месяца это приводит к

ИЗМЕНЕНИЕ ВЫХОДНОГО СИГНАЛА, %

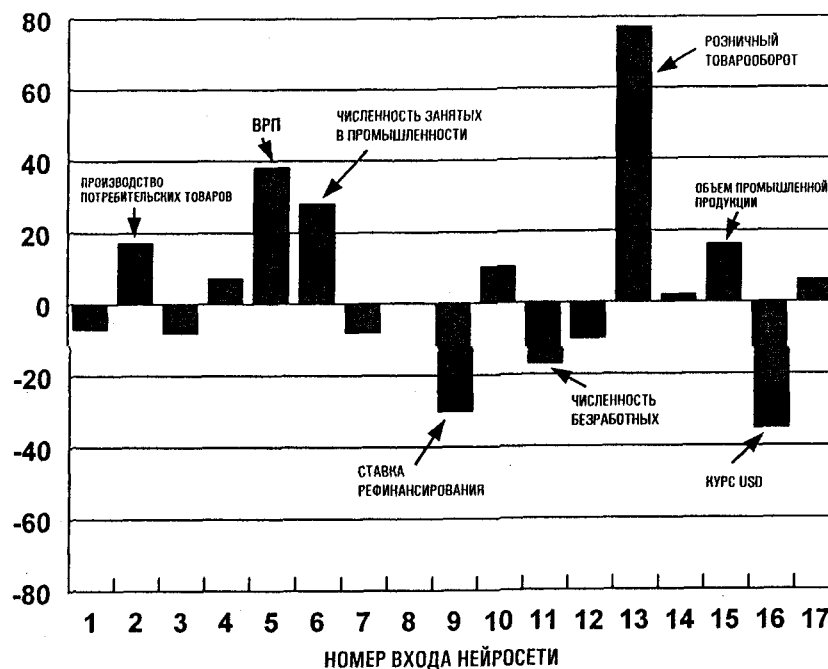


Рис.4. Диаграмма, показывающая степень влияния 50%-ого прироста каждого из входных макропоказателей на величину поступления налогов. Переводя это изменение в относительную шкалу (в % от фактического значения выходного сигнала для данного месяца) и усредняя его по всем месяцам, формируем итоговую таблицу, показывающую среднюю степень влияния изменения каждого из входных макропоказателей на выходную величину.

На рис. 4 эти результаты приведены в графическом виде. Каждый вход нейронной сети пронумерован от 1 до 17, и эта нумерация соответствует рис 2. Диаграмму следует трактовать следующим образом (коснемся только самых существенных макропоказателей).

- Если обеспечить прирост розничного товарооборота на 50 % , то даже без воздействия на другие макропоказатели, можно увеличить поступление налогов в среднем на 80 %.
- Если обеспечить прирост валового регионального продукта на 50 % , то даже без воздействия на другие макропоказатели, можно увеличить поступление налогов в среднем на 40 %.
- Если увеличить численность занятых в промышленности на 50 % , то даже без воздействия на другие макропоказатели, можно увеличить поступление налогов в среднем на 30 %.
- Если обеспечить прирост объема промышленной продукции или производства потребительских товаров на 50 % , то даже без воздействия на другие макропоказатели, можно увеличить поступление налогов почти на 20 %.
- Если вдруг ставка рефинансирования ЦБ РФ увеличится на 50 % , то это приведет к уменьшению поступления налогов в бюджетную систему области на 30% . Это же заключение в инвертированном виде - если ставка рефинансирования снизится на 50 % (т.е. будет не 42 % как в январе 1997 г. , а 21 % годовых), то это приведет к увеличению поступления налогов на 30%.

• Если обменный курс американского доллара увеличится на 50 %, то это приведет к уменьшению поступления налогов на 35 %. В инвертированном виде - снижение курса USD на 50 % может привести к увеличению поступления налогов на 35 %.

Отметим, что некоторые макропоказатели оказались существенными, однако влиять на них на уровне регионального руководства не представляется возможным. К ним относятся ставка рефинансирования ЦБ РФ и обменный курс американского доллара. А вот на производство потребительских товаров, валовой региональный продукт, численность занятых в промышленности, численность безработных, розничный товарооборот и объем промышленной продукции оказывать влияние путем проведения целенаправленной региональной экономической политики в какой-то мере возможно.

Понятно, что подобный анализ достаточно искусственен, поскольку в реальной жизни так не бывает, что можно дать приращение только одному из макропоказателей и считать, что остальные при этом не изменятся. (компоненты входного вектора не являются независимыми). Однако как показали наши дополнительные расчеты, это не меняет кардинальным образом ситуацию. Чувствительные входы нейросети, оказывающие существенное влияние на выходной сигнал, остаются чувствительными, и происходит только некоторое перераспределение их по местам на “пьедестале почета”.

Разработанный инструментарий, с целью его дополнительной проверки, использовался для анализа другого набора статистических данных из области социологии. В разработанном программном обеспечении абсолютно ничего не менялось, и оно было применено к анализу таких данных, где генерируемые гипотезы могут легко проверяться с позиций простой житейской логики. Результаты приведены в приложении.

**Таким образом, на основе проведенных расчетов, можно сделать следующие выводы, которые можно рассматривать как практические рекомендации для того, чтобы увеличить поступление налогов и других обязательных платежей в бюджетную систему.**

- 1. Следует предпринять все меры к тому, чтобы розничный товарооборот увеличивался. Надо сделать так, чтобы людям было выгодно тратить свои деньги на покупки внутри области.**
- 2. Следует предпринимать меры, приводящие к росту валового регионального продукта.**
- 3. Следует стремиться к созданию дополнительных рабочих мест в промышленности.**
- 4. Нельзя допускать спада объема промышленной продукции, производства потребительских товаров и объема платных услуг.**

Если эти выводы представляются очевидными и понятными, и для формулирования этих выводов кому-то не требуется привлечение математических расчетов, то приведенные расчетные данные можно использовать хотя бы как объективное их подтверждение на основе “сухой” и беспристрастной статистики.



## Литература.

1. *Прокофьев А.Е.* “Нейросетевая модель для прогнозных расчетов макроэкономических показателей”. - Вологда : ВНКЦ РАН, 1996. Препринт.
2. *Прокофьев А.Е.* “Результаты экспериментального прогноза основных макроэкономических показателей для Вологодской области в сентябре-декабре 1996 г. и экстраполяционный вариант прогноза социально-экономического развития региона в 1997 г.”. - Вологда : ВНКЦ РАН, 1996. Информационно-аналитическая записка.
3. *Lawrence J.* Introduction to Neural Networks. - Nevada City, CA: CSS, 1994.
4. *Rumelhart D.E., Hinton G.E., Williams R.J.* Learning representations by back-propagation errors // Nature. 1986. Vol. 323. P.533

## **ПРИЛОЖЕНИЕ**

**Постановка задачи.**

В феврале 1997 г. в разных точках Вологодской области силами ВНКЦ РАН был проведен очередной опрос общественного мнения. Всего опрошено 1500 респондентов. Опросная анкета содержала большое количество разнообразных вопросов, среди которых один формулировался следующим образом.

“Оцените потребность вашей семьи в продуктах питания” :

	Товары	Обеспечены в достаточной мере	Нуждаемся, но нет средств	Нет потребности	Затрудняюсь ответить
1	Сметана	√			
2	Мясо		√		
3	Рыба	√			
4	Масло животное		√		
5	Фрукты		√		
6	Картофель	√			
7	Кондитерские изделия		√		
8	Табак				√
9	Винно-водочные изделия			√	
10	Хлебобулочные изделия	√			

Здесь для примера символом √ отмечены ответы одного из респондентов

Все анкеты случайным образом были разложены в 15 стопок по 100 штук в каждой. Каждую стопку будем называть группой. В результате обработки анкет по каждой группе была получена следующая таблица:

**Таблица 1**

группы	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41
1	42	46	4	7	44	47	3	6	31	53	6	9	40	49	6	5	17	68	8	6	88	4	2	6	22	62	5	11	35	17	34	14	16	23	41	20	76	12	1	10	51
2	40	43	6	10	46	41	4	9	31	56	4	9	46	41	4	9	24	62	5	8	84	4	7	5	22	57	7	14	21	13	52	14	20	10	47	23	85	9	0	6	37
3	47	43	5	5	42	48	3	6	28	59	4	8	48	43	1	7	12	74	4	8	88	3	4	5	18	58	12	11	31	12	46	11	16	17	58	9	81	12	0	7	51
4	44	46	4	6	44	46	2	7	37	51	7	5	39	52	6	3	24	69	2	4	85	7	5	3	25	66	5	4	32	23	38	6	15	24	44	16	80	17	1	2	45
5	32	54	5	8	35	52	5	8	28	65	0	7	42	47	4	7	13	81	0	6	81	9	3	7	20	68	4	8	23	23	39	25	12	19	50	19	73	20	1	6	48
6	31	48	3	17	36	51	2	11	30	58	2	10	39	46	4	11	16	68	4	12	75	7	7	11	18	68	4	10	25	16	44	15	17	21	43	19	76	12	0	12	40
7	41	52	4	3	50	41	2	7	31	58	2	8	39	49	6	6	21	68	6	5	88	4	6	2	30	57	7	6	29	22	40	9	23	19	40	18	76	19	0	5	46
8	38	53	1	8	37	52	3	7	26	66	2	6	34	54	5	7	18	65	8	9	87	5	2	5	16	60	11	13	25	21	38	16	14	17	46	22	80	11	2	7	46
9	38	45	7	9	41	47	3	9	31	56	6	7	41	46	5	8	17	68	5	10	79	8	4	8	19	59	11	9	28	20	42	9	14	24	44	16	73	15	5	7	53
10	45	39	8	8	43	47	2	7	36	54	5	5	44	41	5	10	25	66	4	5	85	7	4	4	24	60	8	8	28	22	36	12	19	22	39	18	76	16	0	8	41
11	42	44	2	12	45	47	0	8	40	50	3	7	41	47	2	10	20	68	3	8	82	3	4	10	24	61	7	7	30	19	38	12	17	22	47	13	75	17	0	8	49
12	43	48	4	4	48	46	1	4	30	62	2	4	46	47	3	3	13	72	4	8	88	3	4	5	21	63	8	8	29	18	46	7	19	14	51	16	79	15	1	5	39
13	38	54	2	6	44	50	3	3	32	53	4	11	40	49	1	9	15	76	2	7	86	4	4	5	24	64	3	9	20	24	45	11	10	20	52	18	79	12	1	7	46
14	42	48	1	9	34	60	0	5	29	62	2	7	43	47	0	10	17	73	0	9	78	9	3	10	22	59	4	15	27	20	37	15	16	20	43	20	72	19	1	8	47
15	44	46	3	5	34	56	5	5	39	47	5	9	45	46	5	4	21	66	5	7	84	8	3	4	24	57	7	11	26	18	45	9	16	21	47	13	77	14	4	5	41

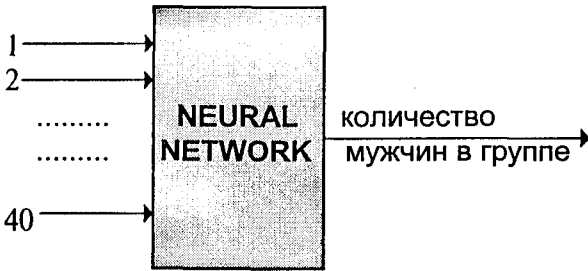
**Кодировка Таблицы 1**

- колонка 1 - количество людей на вопрос о СМЕТАНЕ, ответивших так : обеспечены в достаточной мере
- колонка 2 - количество людей на вопрос о СМЕТАНЕ, ответивших так : нуждаемся, но нет средств
- колонка 3 - количество людей на вопрос о СМЕТАНЕ, ответивших так : нет потребности
- колонка 4 - количество людей на вопрос о СМЕТАНЕ, ответивших так : затрудняюсь ответить
- колонка 5 - количество людей на вопрос о МЯСЕ, ответивших так : обеспечены в достаточной мере
- колонка 6 - количество людей на вопрос о МЯСЕ, ответивших так : нуждаемся, но нет средств
- колонка 7 - количество людей на вопрос о МЯСЕ, ответивших так : нет потребности
- колонка 8 - количество людей на вопрос о МЯСЕ, ответивших так : затрудняюсь ответить
- колонка 9 - количество людей на вопрос о РЫБЕ, ответивших так : обеспечены в достаточной мере
- колонка 10 - количество людей на вопрос о РЫБЕ, ответивших так : нуждаемся, но нет средств
- колонка 11 - количество людей на вопрос о РЫБЕ, ответивших так : нет потребности
- колонка 12 - количество людей на вопрос о РЫБЕ, ответивших так : затрудняюсь ответить
- колонка 13 - количество людей на вопрос о МАСЛЕ, ответивших так : обеспечены в достаточной мере
- колонка 14 - количество людей на вопрос о МАСЛЕ, ответивших так : нуждаемся, но нет средств
- колонка 15 - количество людей на вопрос о МАСЛЕ, ответивших так : нет потребности
- колонка 16 - количество людей на вопрос о МАСЛЕ, ответивших так : затрудняюсь ответить
- колонка 17 - количество людей на вопрос о ФРУКТАХ, ответивших так : обеспечены в достаточной мере
- колонка 18 - количество людей на вопрос о ФРУКТАХ, ответивших так : нуждаемся, но нет средств
- колонка 19 - количество людей на вопрос о ФРУКТАХ, ответивших так : нет потребности
- колонка 20 - количество людей на вопрос о ФРУКТАХ, ответивших так : затрудняюсь ответить
- колонка 21 - количество людей на вопрос о КАРТОФЕЛЕ, ответивших так : обеспечены в достаточной мере
- колонка 22 - количество людей на вопрос о КАРТОФЕЛЕ, ответивших так : нуждаемся, но нет средств
- колонка 23 - количество людей на вопрос о КАРТОФЕЛЕ, ответивших так : нет потребности
- колонка 24 - количество людей на вопрос о КАРТОФЕЛЕ, ответивших так : затрудняюсь ответить
- колонка 25 - количество людей на вопрос о КОНФЕТАХ, ответивших так : обеспечены в достаточной мере
- колонка 26 - количество людей на вопрос о КОНФЕТАХ, ответивших так : нуждаемся, но нет средств
- колонка 27 - количество людей на вопрос о КОНФЕТАХ, ответивших так : нет потребности
- колонка 28 - количество людей на вопрос о КОНФЕТАХ, ответивших так : затрудняюсь ответить
- колонка 29 - количество людей на вопрос о ТАБАКЕ, ответивших так : обеспечены в достаточной мере
- колонка 30 - количество людей на вопрос о ТАБАКЕ, ответивших так : нуждаемся, но нет средств
- колонка 31 - количество людей на вопрос о ТАБАКЕ, ответивших так : нет потребности
- колонка 32 - количество людей на вопрос о ТАБАКЕ, ответивших так : затрудняюсь ответить
- колонка 33 - количество людей на вопрос о ВИНЕ, ответивших так : обеспечены в достаточной мере
- колонка 34 - количество людей на вопрос о ВИНЕ, ответивших так : нуждаемся, но нет средств
- колонка 35 - количество людей на вопрос о ВИНЕ, ответивших так : нет потребности
- колонка 36 - количество людей на вопрос о ВИНЕ, ответивших так : затрудняюсь ответить
- колонка 37 - количество людей на вопрос о ХЛЕБЕ, ответивших так : обеспечены в достаточной мере
- колонка 38 - количество людей на вопрос о ХЛЕБЕ, ответивших так : нуждаемся, но нет средств
- колонка 39 - количество людей на вопрос о ХЛЕБЕ, ответивших так : нет потребности
- колонка 40 - количество людей на вопрос о ХЛЕБЕ, ответивших так : затрудняюсь ответить
- колонка 41 - количество мужчин в группе

**Требуется установить :** численные изменения в каких колонках оказывают существенное влияние на количество мужчин в группе.

Другими словами - какие варианты ответов *существенным образом характеризуют мужчину* (в пределах рассматриваемой выборки респондентов и рассматриваемых признаков).

Задача сводится к построению нейронной сети прямого распространения, имеющей 40 входов ( данные из колонок 1-40 ) и один выход (данные из колонки 41), обучению этой сети на примерах (строки с 1 по 15 из таблицы 1), и исследованию чувствительности каждого входа. Схема нейронной сети :



**Кодировка входов нейронной сети**

- 1 - количество людей в группе на вопрос о СМЕТАНЕ, ответивших так : обеспечены в достаточной мере
- 2 - количество людей в группе на вопрос о КАРТОФЕЛЕ, ответивших так : нет потребности
- 3 - количество людей в группе на вопрос о РЫБЕ, ответивших так : обеспечены в достаточной мере
- 4 - количество людей в группе на вопрос о ВИНЕ, ответивших так : обеспечены в достаточной мере
- 5 - количество людей в группе на вопрос о КАРТОФЕЛЕ, ответивших так : затрудняюсь ответить
- 6 - количество людей в группе на вопрос о ФРУКТАХ, ответивших так : затрудняюсь ответить
- 7 - количество людей в группе на вопрос о МАСЛЕ, ответивших так : нет потребности
- 8 - количество людей в группе на вопрос о ВИНЕ, ответивших так : затрудняюсь ответить
- 9 - количество людей в группе на вопрос о ТАБАКЕ, ответивших так : нуждаемся но нет средств
- 10 - количество людей в группе на вопрос о МЯСЕ, ответивших так : затрудняюсь ответить
- 11 - количество людей в группе на вопрос о ВИНЕ, ответивших так : нуждаемся но нет средств
- 12 - количество людей в группе на вопрос о ТАБАКЕ, ответивших так : обеспечены в достаточной мере
- 13 - количество людей в группе на вопрос о КАРТОФЕЛЕ, ответивших так : обеспечены в достаточной мере
- 14 - количество людей в группе на вопрос о МАСЛЕ, ответивших так : обеспечены в достаточной мере
- 15 - количество людей в группе на вопрос о СМЕТАНЕ, ответивших так : нуждаемся но нет средств
- 16 - количество людей в группе на вопрос о МЯСЕ, ответивших так : обеспечены в достаточной мере
- 17 - количество людей в группе на вопрос о МЯСЕ, ответивших так : нет потребности
- 18 - количество людей в группе на вопрос о ХЛЕБЕ, ответивших так : затрудняюсь ответить
- 19 - количество людей в группе на вопрос о ФРУКТАХ, ответивших так : нуждаемся но нет средств
- 20 - количество людей в группе на вопрос о МЯСЕ, ответивших так : нуждаемся но нет средств
- 21 - количество людей в группе на вопрос о МАСЛЕ, ответивших так : затрудняюсь ответить
- 22 - количество людей в группе на вопрос о КАРТОФЕЛЕ, ответивших так : нуждаемся но нет средств
- 23 - количество людей в группе на вопрос о КОНФЕТАХ, ответивших так : обеспечены в достаточной мере
- 24 - количество людей в группе на вопрос о КОНФЕТАХ, ответивших так : нет потребности
- 25 - количество людей в группе на вопрос о ФРУКТАХ, ответивших так : нет потребности
- 26 - количество людей в группе на вопрос о СМЕТАНЕ, ответивших так : нет потребности
- 27 - количество людей в группе на вопрос о ТАБАКЕ, ответивших так : затрудняюсь ответить
- 28 - количество людей в группе на вопрос о ВИНЕ, ответивших так : нет потребности
- 29 - количество людей в группе на вопрос о ХЛЕБЕ, ответивших так : нуждаемся но нет средств
- 30 - количество людей в группе на вопрос о ХЛЕБЕ, ответивших так : обеспечены в достаточной мере
- 31 - количество людей в группе на вопрос о РЫБЕ, ответивших так : нуждаемся но нет средств
- 32 - количество людей в группе на вопрос о ТАБАКЕ, ответивших так : нет потребности
- 33 - количество людей в группе на вопрос о РЫБЕ, ответивших так : нет потребности
- 34 - количество людей в группе на вопрос о РЫБЕ, ответивших так : затрудняюсь ответить
- 35 - количество людей в группе на вопрос о КОНФЕТАХ, ответивших так : затрудняюсь ответить
- 36 - количество людей в группе на вопрос о МАСЛЕ, ответивших так : нуждаемся но нет средств
- 37 - количество людей в группе на вопрос о СМЕТАНЕ, ответивших так : затрудняюсь ответить
- 38 - количество людей в группе на вопрос о ХЛЕБЕ, ответивших так : нет потребности
- 39 - количество людей в группе на вопрос о ФРУКТАХ, ответивших так : обеспечены в достаточной мере
- 40 - количество людей в группе на вопрос о КОНФЕТАХ, ответивших так : нуждаемся но нет средств

На рис. 1 графически показаны результаты расчетов. Оставляем их без комментариев. Напомним лишь одну известную русскую народную поговорку : “У кого что болит, тот о том и говорит”. Интересно, что нейросеть выделила один необычный фактор, не приходящий на ум человеку-эксперту изначально (в этом ее достоинство). Речь идет о фруктах. Один человек дал такое объяснение : “Мужчинам нужны фрукты, чтобы заглаживать свои провинности перед женщинами и детьми”.

### ИЗМЕНЕНИЕ ВЫХОДНОГО СИГНАЛА, %

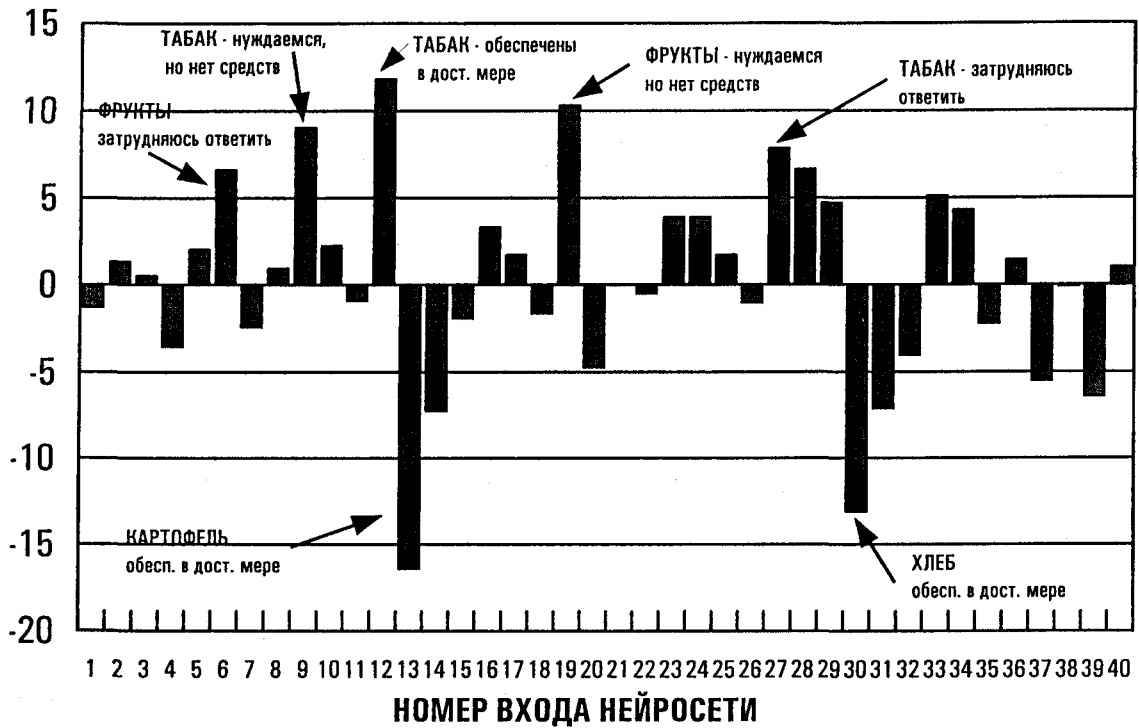


Рис. 1. Чувствительность входов нейронной сети. МУЖЧИНЫ.

На диаграмме 2 приведены аналогичные результаты показывающие существенные факторы для женщин. Они укладываются в формулу "Чем я завтра буду кормить свою семью?" Таковы суровые реалии жизни женщин-северянок.

### ИЗМЕНЕНИЕ ВЫХОДНОГО СИГНАЛА, %

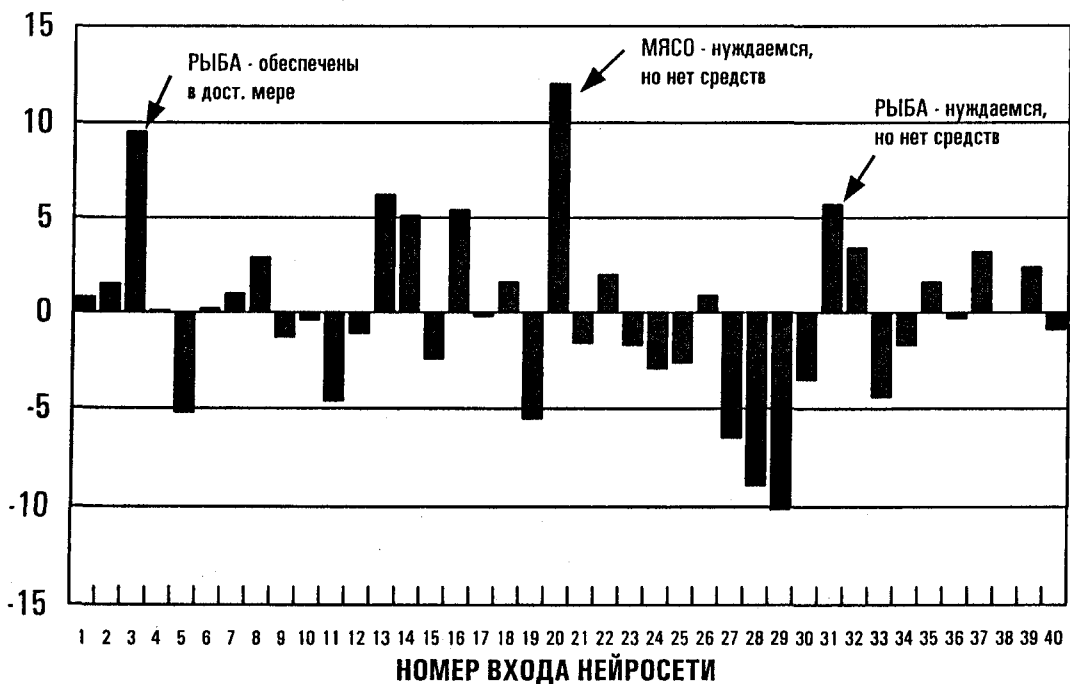


Рис. 2 Чувствительность входов нейронной сети. ЖЕНЩИНЫ.

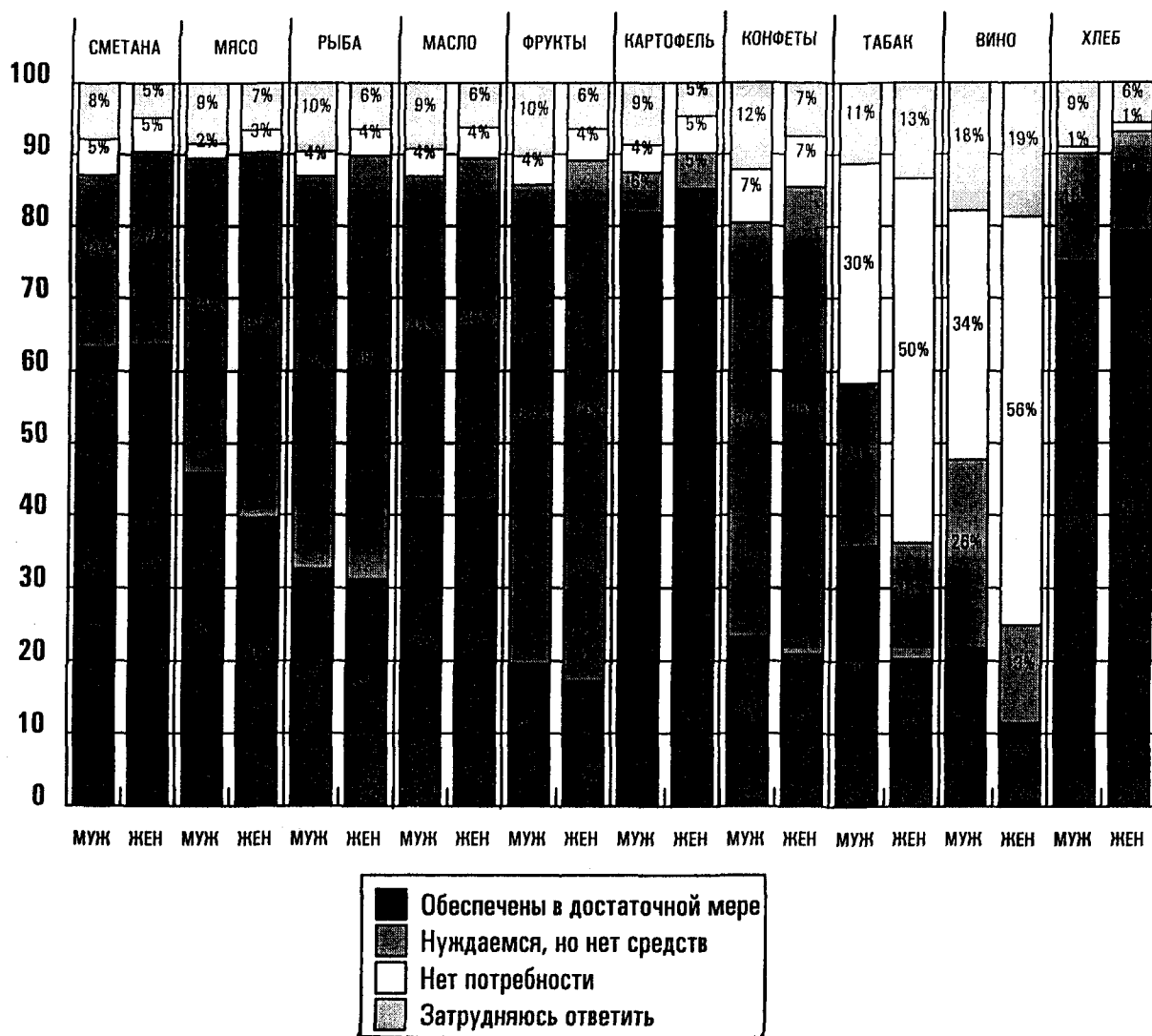


Рис. 3. Сравнение структуры ответов для мужчин и женщин. Традиционный расчет .

На рис. 3 показана структура ответов для мужчин и женщин. Приведенные результаты показывают, например, что в вопросе о сметане :

- 63 % всех опрошенных мужчин ответили - обеспечены в достаточной мере
- 64 % всех опрошенных женщин ответили - обеспечены в достаточной мере
- 24 % всех опрошенных мужчин ответили - нуждаемся, но нет средств
- 27 % всех опрошенных женщин ответили - нуждаемся, но нет средств
- 5 % всех опрошенных мужчин ответили - нет потребности
- 5 % всех опрошенных женщин ответили - нет потребности
- 8 % всех опрошенных мужчин ответили - затрудняюсь ответить
- 5 % всех опрошенных женщин ответили - затрудняюсь ответить

И так далее по остальным продуктам питания. Некоторое различие имеется в отношении к вину и табаку, что естественно.

Таким образом, с точки зрения традиционного статистического анализа особой разницы между мужчинами и женщинами по основным продуктам питания нет, и сделать выводы аналогичные рис. 1 - 2 затруднительно.