

ОПРЕДЕЛЕНИЕ ЗНАЧАЩИХ ФАКТОРОВ ПРИ РЕШЕНИИ ЗАДАЧИ ПРОГНОЗИРОВАНИЯ СТРАХОВЫХ ВЗНОСОВ И СТРАХОВЫХ ПЛАТЕЖЕЙ В УСЛОВИЯХ СТРАХОВОЙ КОМПАНИИ

Яцук Д.М., Светличная В.А.

Донецкий национальный технический университет

Страховая компания занимается страхованием имущества (движимого или недвижимого) или жизни человека. Основным источником поступлений страховой компании являются взносы страховых платежей или, говоря проще – страховые взносы, а основным источником расходов являются страховые выплаты.

Суть страхования заключается в следующем:

Лицо (юридическое или физическое) делает в страховую компанию запрос на страхование имущества или жизни. Представитель страховой компании агент рассматривает поступившую заявку и принимает по ней решение страховать или нет. По заявке определяются ее основные параметры и в соответствии с действующим законодательством устанавливается сумма страхового взноса. После оплаты страхового взноса лицо получает полис (договор страхования) в котором указаны сроки действия полиса и правила и условия действия данного полиса.

Если в период действия договора не происходит случая, указанного в правилах полиса, то выплат не производится, и страховая компания получает только поступления.

Если в период действия договора происходит случай, указанный в правилах полиса, то страхователь сообщает страховой компании об этом случае. После предоставления документов, страхования компания осуществляет выплату страхователю суммы через кассу или денежным переводом на счет.

В процессе функционирования страховой компании, она имеет право полученный средства положить на депозитный счет. В условиях рыночных отношений каждое предприятие стремится получить большую прибыль, то есть страховая компания желает как можно больше средств положить на депозитный счет, с целью получения большей прибыли. Но поскольку страховая компания обязана в надлежащем порядке производить выплаты, то возникает задача определения оптимального баланса между средствами, которые должны быть доступны, и средствами, используемыми для депозитного вложения. Для определения этого баланса необходимо использовать прогнозирование. В задаче прогнозирования страховых взносов и выплат рассматривалось страхование автотранспорта как имущества (одного из основных видов страхования).

При решении задачи были выделены факторы, которые влияют на прогноз: время (год и месяц); место положение (страна, область, город); населенность города; средняя населенность на квадратный метр; сезонность (время года: весна, лето, осень, зима); минимальная заработная плата; средняя заработная плата населения; среднее количество проданных автомобилей за месяц; среднее количество аварий за месяц; количество страховых случаев; средний стаж водителей; средняя стоимость приобретаемых автомобилей; процент кредита в банках; затраты на рекламу; уровень инфляции; стоимость международной валюты; туристический бизнес (основные потоки туристов на въезд и выезд из страны); инвестиции зарубежных инвесторов; основные течения рынков; параметры грунта; состояние дорог; среднее расстояние поездок; стоимость основных источников топлива (ДТ, А92, А95 и другие); сезон отпусков; количество угонов; количество испорченных автомобилей в результате воровства;

количество несчастных случаев с автомобилем без участия человека (упало дерево, ударила молния и прочее).

Для задачи прогнозирования такое количество факторов является большим. Также в результате исследования (тестирования) и анализа мною было выявлено, что некоторые факторы являются взаимосвязанными прямой зависимостью (рост значения одного фактора сопровождается ростом значения другого, аналогично со снижением значений факторов) и обратной (рост значения одного фактора сопровождается снижением значения другого, аналогично со снижением).

Таким образом возникла задача определения значащих факторов, влияющих на прогноз.

Были рассмотрены такие методы отбора факторов как:

Метод прямого отбора (Forward Selection)

Процедуру отбора независимых переменных по методу Forward Selection можно представить в виде блок-схемы [1], изображенной на рис. 1.

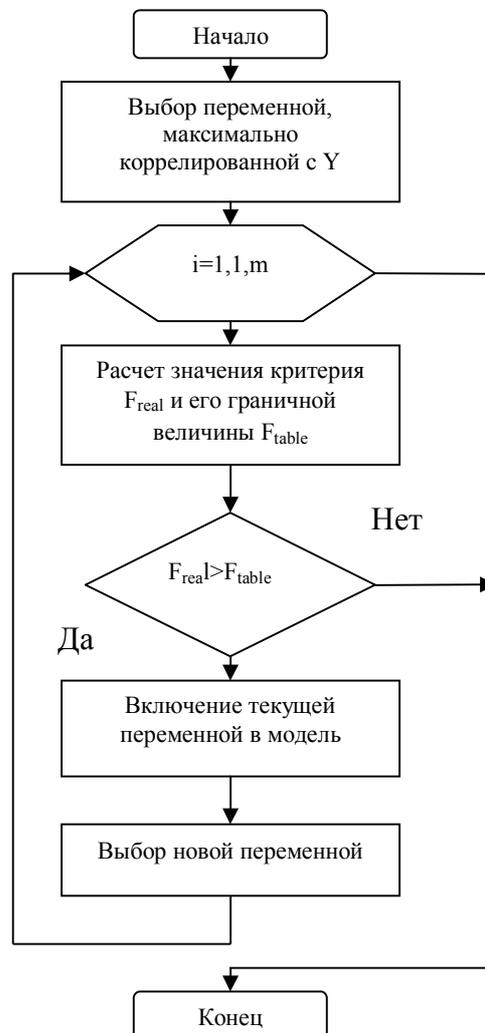


Рисунок 1 - Блок-схема метода Forward Selection

Метод обратного исключения (Backward Elimination)

На рис. 2 приведена блок-схема процедуры отбора факторов методом обратного исключения [1].

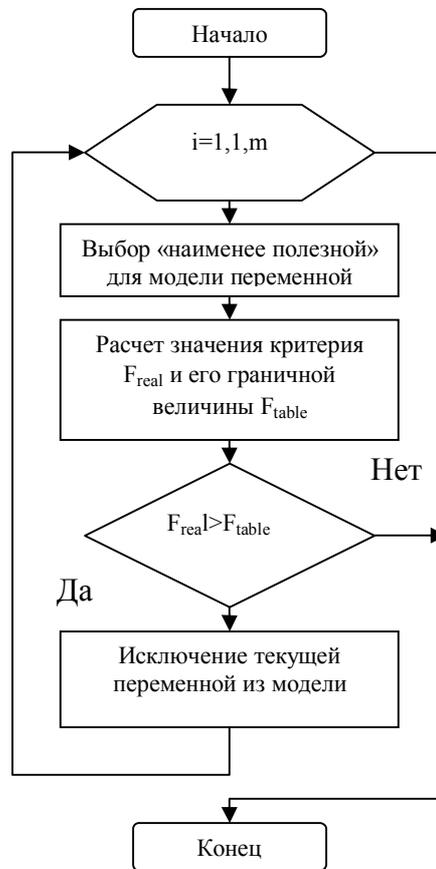


Рисунок 2 – Блок-схема метода Backward Elimination

Метод последовательного отбора (Stepwise)

Ниже (рис. 3) приведена блок-схема соответствующего метода [1].

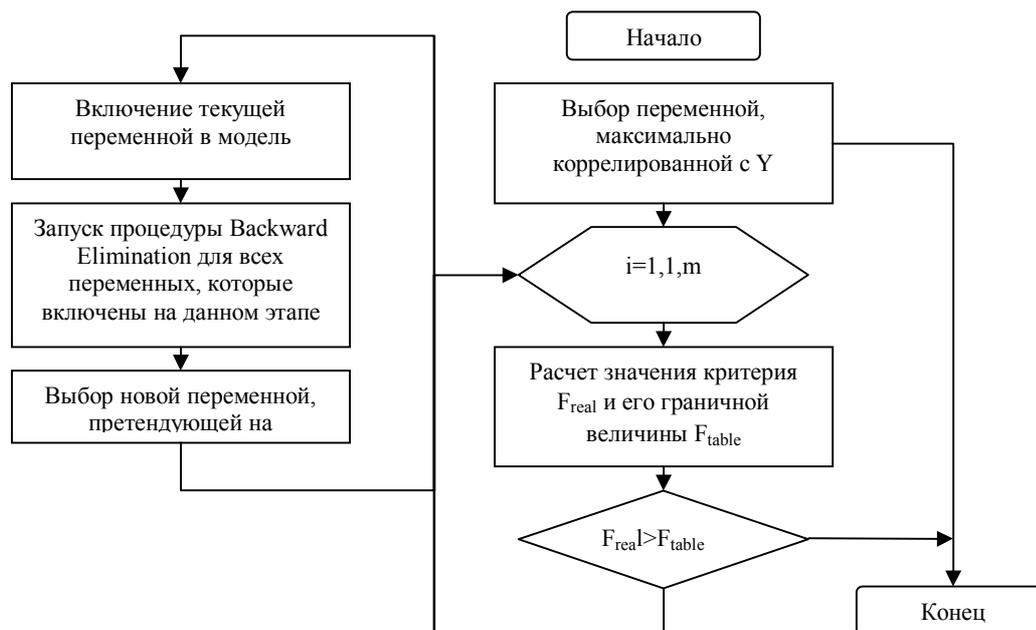


Рисунок 3 – Блок-схема метода Stepwise

Отбор на основе «лучших подмножеств» (Best Subsets)

Данный метод предусматривает, что аналитик изначально должен определиться с максимально возможным количеством P входных признаков, задействованных в модели. Далее поступают следующим образом: по всему перечню переменных строятся P групп моделей, причем первую такую группу составляют все однофакторные модели, которые только можно построить на имеющемся наборе признаков, соответственно, во вторую попадают все двухфакторные модели и т. д. вплоть до P -й группы (куда входят все модели, включающие p факторов). По всем полученным моделям рассчитываются статистические показатели, характеризующие качество модели, такие как коэффициент детерминации R^2 , скорректированный коэффициент детерминации R_{adj}^2 , стандартная ошибка оценивания s (равная квадратному корню из суммы квадратов ошибки модели, приходящейся на одну степень свободы MSE) и другие. Затем из каждой полученной группы отбираются только k самых лучших моделей с точки зрения рассчитанных показателей, после чего формируется отчет, включающий K и P моделей, на основании которого аналитик непосредственно делает окончательное заключение [2].

Для уменьшения количества факторов использовался метод на основе «лучших подмножеств». Были взяты ежемесячные статистические данные с января 2004 года по декабрь 2008 года, и составили 60 значений по каждому фактору. Была проанализирована динамика изменения факторов и отмечена зависимость одних факторов от других. Проведя анализ были выделены 10 факторов: время (год, месяц); количество страховых случаев; средняя зарплата; стаж водителей; средняя стоимость приобретаемых автомобилей; средний курс доллара; уровень инфляции; стоимость международной валюты; туристический бизнес (основные потоки туристов на въезд и выезд из страны); инвестиции зарубежных инвесторов, остальные отброшены как малозначимые, то есть их влияние на прогноз оценено менее чем 0,001 и считается мало значимым. Каждый из десяти факторов был проанализирован на зависимость с оставшимися девятью факторами методом на основе «лучших подмножеств». В результате получилось сокращение до 6 факторов:

1. время (год, месяц);
2. количество страховых случаев;
3. средняя зарплата;
4. стаж водителей;
5. средняя стоимость приобретаемых автомобилей;
6. средний курс доллара.

После определения основных факторов путем исследования (тестирования) и анализа производился поиск метода прогноза. Были проанализированы и протестированы такие методы прогноза:

1. линейная регрессионная модель;
2. логистическая регрессионная модель;
3. модель авторегрессии;
4. скользящего среднего;
5. метод экстраполяции;
6. метод расстановки;
7. и другие.

Все эти методы были проанализированы на основе статистических данных. В результате каждый из методов давал ошибку прогноза не менее 0,1. А некоторые методы требовали линейной зависимости или трудоемкого процесса вычисления коэффициентов.

Однако, метод прогноза с использованием нейронных сетей позволил устранить эти недостатки. В результате тестирования была спроектирована нейронная сеть. После детального анализа была выделена такая структура нейронной сети:

Таблица 1

Структура нейронной сети

№	Функция	Скрытые слои	Число нейронов	Функции активации	Метод обучения	Циклы	Ошибка
1	newff()	2	40-20-10-1	2-2-1-1	trainrp	1000	10-3

Выводы

1. Все факторы, влияющие на прогноз были исследованы и проанализированы;
2. Был произведен отбор факторов методом на основе «лучших подмножеств»;
3. Среди всех факторов были выделены основные значащие, остальные – имеют зависимость от основных значащих;
4. Были проанализированы методы прогноза, выявлены их недостатки и выбран метод прогноза наиболее соответствующий поставленной задаче.

Литература

- [1] Эконометрика: Учебник / Под ред И. И. Елисейевой - М.: Финансы и статистика, 2005. Электронный ресурс. Режим доступа: http://www.basegroup.ru/library/analysis/regression/feature_selection/
- [2] Ханк Д. Э., Райтс А. Дж., Уичерн Д. У.. Бизнес-прогнозирование. 7-е изд. / Пер. с англ. - М.: Издательский дом "Вильямс", 2003. Электронный ресурс: http://www.basegroup.ru/library/analysis/regression/feature_selection/