

Реализация принятия решения при выборе мотоцикла для комфортного вождения с помощью байесовских сетей доверия в системе Netica

Худяков Алексей Викторович

Приамурский государственный университет имени Шолом-Алейхема

Студент

Научный руководитель:

Баженов Руслан Иванович

Приамурский государственный университет имени Шолом-Алейхема

к.п.н., доцент, зав. кафедрой информационных систем, математики и правовой информатики

Аннотация

В работе рассматривается задача по выбору мотоцикла для комфортного вождения, учитывая параметры заказчика. Исследование проводится в системе создания байесовских сетей доверия Netica. Описан алгоритм создания байесовской сети доверия.

Ключевые слова: Байесовские сети доверия, система Netica, оптимальный выбор.

Implementation of decision-making when choosing a motorcycle for comfortable driving using Bayesian trust networks in the Netica system

Hudyakov Aleksey Viktorovich

Sholom-Aleichem Priamursky State University

Student

Research Supervisor:

Bazhenov Ruslan Ivanovich

Sholom-Aleichem Priamursky State University

Candidate of pedagogical sciences, associate professor, Head of the Department of Information Systems, Mathematics and Legal Informatics

Abstract

The paper considers the problem of choosing a motorcycle for comfortable driving, taking into account the parameters of the customer. The research is carried out in the system for creating Bayesian trust networks Netica. An algorithm for creating a Bayesian trust network is described

Keywords: Bayesian belief networks, system Netica, optimum choice.

При выборе мотоцикла люди часто задаются вопросом, какой мотоцикл себе выбрать для комфортной езды, какой тип посадки им подойдет больше

или какой будет лучше по ключевым параметрам. В подобных случаях помогает вычислительная модель под названием байесовские сети доверия, которая позволяет существенно сократить время для принятия верного решения, исключив возможность ошибки и повторения ранее проведённых этапов анализа ситуации для правильного выбора.

Данная модель дает возможность определить связи между переменными и вычислить их вероятностные значения, учитывая, что события условно независимы. Это дает возможность выстроить сеть без построения больших таблиц вероятностей, вследствие чего уменьшается количество и объем вычислений.

Психологические эксперименты показали, что люди часто неверно оценивают вероятность события, на основе полученного опыта (апостериорная вероятность), поскольку игнорируют саму вероятность предположения (априорная вероятность). Поэтому правильный результат по формуле Байеса может сильно отличаться от интуитивно ожидаемого. Теорема Байеса названа в честь её автора Томаса Байеса (1702–1761) – английского математика и священника, который первым предложил использование теоремы для корректировки убеждений, основываясь на 11 обновлённых данных. Его работа «An Essay towards solving a Problem in the Doctrine of Chances» впервые опубликована в 1763 году, через 2 года после смерти автора. До того, как посмертная работа Байеса была принята и прочитана в Королевском обществе, она была значительно отредактирована и обновлена Ричардом Прайсом. Однако эти идеи не предавались публичной огласке до тех пор, пока не были вновь открыты и развиты Лапласом, впервые опубликовавшим современную формулировку теоремы в своей книге 1812 года «Аналитическая теория вероятностей».

В статье А.В.Сироткин описал алгоритмы построение алгебраической байесовской сети, семантически эквивалентной многосвязной байесовской сети доверия [3]. Г.Г.Маклакова оценила качество услуг телекоммуникационной сети дистанционного обучения на основе байесовских сетей доверия [4]. Д.А. Яковлев и И.С. Синева изучили процесс детектирование веб-роботов, используя в работе байесовские сети [5]. В работе А.Н. Размахниной и Р.И. Баженова был реализован процесс принятия решений с помощью байесовских сетей доверия в системе Hugin [6]. А.Н. Терентьев, П.И. Бидюк и Л.А. Коршевнюк предложили новый, более простой и точный алгоритм вероятностного вывода в байесовских сетях на основе обучающих данных [7]. Так же в сегменте иноязычных научных исследований байесовские сети активно изучаются и применяются для решения задач. Н. Gan, Y. Zhang, Q. Song использовали байесовскую сеть доверия для положительного немеченого изучения с неопределенностью в решении [8]. Авторами T.D. Phan, J.C.R. Smart, S.J.Capon, W.L. Hadwen, O.Sahin был проведен систематический обзор в сфере управления водными ресурсами, используя байесовские сети [9].

Целью исследования является реализация принятия решений при выборе мотоцикла с помощью байесовских сетей доверия в системе Netica.

Для начала рассмотрим интерфейс выбранной программы. Netica — достаточно удобная в работе программа для работы с графовыми вероятностными моделями. Она имеет интуитивно понятный интерфейс пользователя для ввода данных. Для выполнения многих операций достаточно воспользоваться несколькими щелчками мыши, что облегчает работу пользователя в системе. Netica является продуктом компании Norsys Software Corp. Программа имеет коммерческую и бесплатную версию. Для использования бесплатной версии, необходимо скачать приложение на официальном сайте <http://www.norsys.com/>, оставить диалоговое окно для ввода пароля пустым и нажать 'Limited Mode'. Бесплатная демо-версия является полнофункциональной, но ограничивает созданную модель по размеру.

Для решения поставленной задачи необходима таблица, содержащая основные параметры мотоциклов, это будут наши входные данные. Netica предоставляет возможность добавления входных данных, посредством загрузки файла Microsoft Excel. Однако важно учесть, что заголовок столбца не должен начинаться с символов: /, ,, .

Рассмотрим пример задачи. Для выбора мотоцикла, клиенту предоставляется огромный выбор технической составляющей каждой модели мотоцикла. Такие как: Объем двигателя, количество тактов, расход топлива, класс мотоцикла. Цель задачи определить вероятность покупки той или иной модели мотоцикла, учитывая заявленные требования клиента.

Рассмотрим интерфейс программы Netica.

На рисунке 1 указано расположение кнопок, выполняемые ими функции указаны в таблице 1.

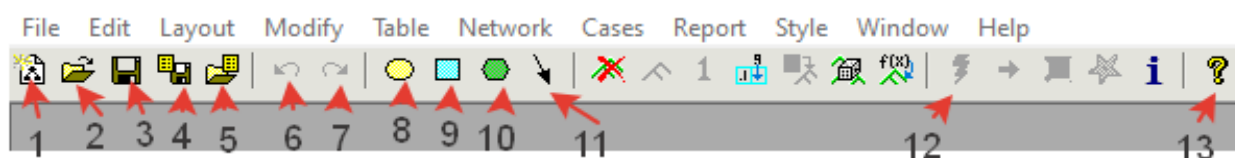


Рис. 1. Панель программы Netica.

Таблица 1. Функции кнопок Netica

№	Наименование	Функция
1	New Net	Создать новую сеть
2	Open	Открыть файл
3	Save	Сохранить
4	Save Case	Сохранить как ...
5	Get Case From File	Добавить таблицу данных
6	Undo	Предыдущее действие
7	Redo	Сделать вновь отмененное действие
8	Add Nature Node	Добавить узел

9	Add Decision Node	Добавить узел с таблицей вероятностей
10	Add Utility Node	Добавить вспомогательный узел
11	Add Link	Добавить связь
12	Compile	Компилировать сеть
13	Help Topics	Помощь

Составим таблицу с входными данными для программы Netica. Таблица заполняется данными для будущей модели системы. Для получения данных были посещены интернет площадки по продаже и покупке мотоциклов. Такие как drom.ru avito.ru Выбраны варианты с разными классами, например: спорт, питбайк, эндуро. В таблице (рис.2) первые строки служат для названия узлов в программе, строки 2-15 являются вариантами вероятности выбора того или иного узла.

	A	B	C	D	E	F
1	Marka-model-god	Obem dvigatela	kolichestvo taktov	cena	rashod-na-100km	klass
2	Harley-Davidson Sportster Forty-Eight-2013		1200	2	690000	4 choper
3	BSE Z5-2020		250	4	169000	3,2 enduro
4	Yamaha YZF1000R-1997		1000,00	4	225000	3,7 sport
5	Kawasaki Ninja 1000 ABS-2011		1000	4	450000	5,6 sport
6	Honda C110 JA07-2011		110,00	4	90000	1,5 pitbaik
7	Honda Dio 110-2014		110,00	4	90000	1,3 pitbaik
8	Honda CBR650F-2013		650,00	4	502000	5 sport
9	Wels MX125-2020		250,00	4	105000	3,4 enduro
10	Yamaha FJR 1300-2001		1300,00	4	345000	4,9 choper
11	Honda Shadow VT1100C-1987		1100,00	4	212000	5,4 choper
12	Orion-2021		49,00	2	28000	1 detskiy
13	Honda DN-01-2008		700,00	4	406000	3,7 choper
14	BMW R1150 GS-2001		1150,00	4	380000	6,1 choper
15	Honda VTR1000-1999		1000,00	4	240000	4,9 sport
16						

Рис. 2. Результат выборки

После заполнения таблицы данными, переходим в программу Netica. Создаем новую сеть: File – New – Network (рис.3).

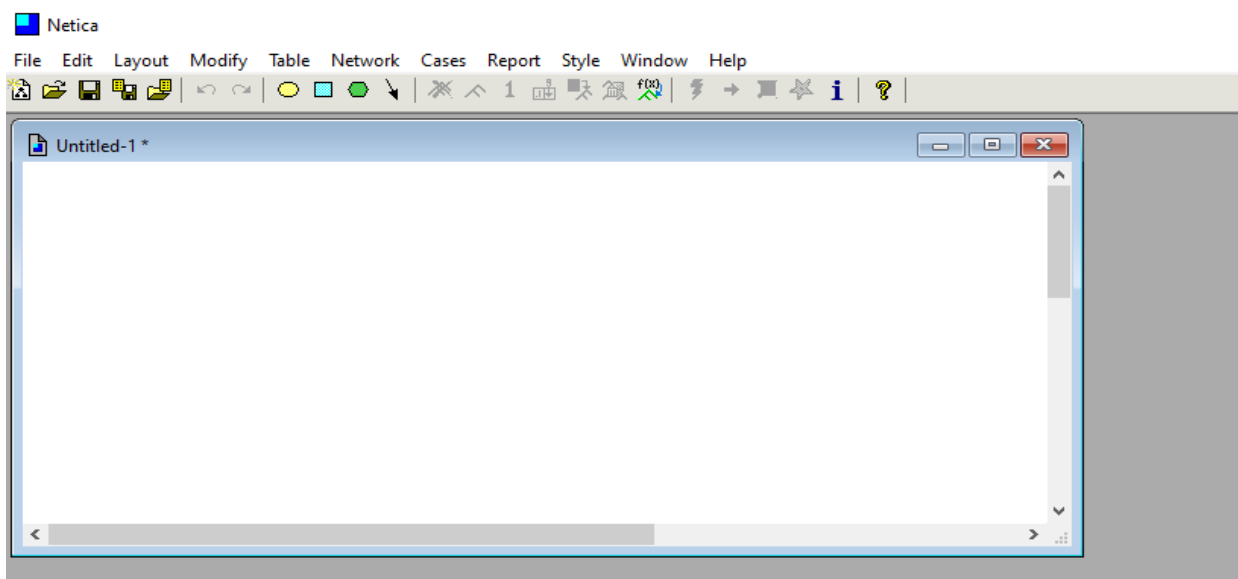


Рис. 3. Создание новой сети

Далее необходимо обучить сеть согласно данным, созданных ранее в MS Excel. Для этого используем команду Cases - Learn - Add Case File Nodes,

в появившемся окне находим наш файл MS Excel и нажимаем "Открыть". Результат данной операции показан на рис. 4.

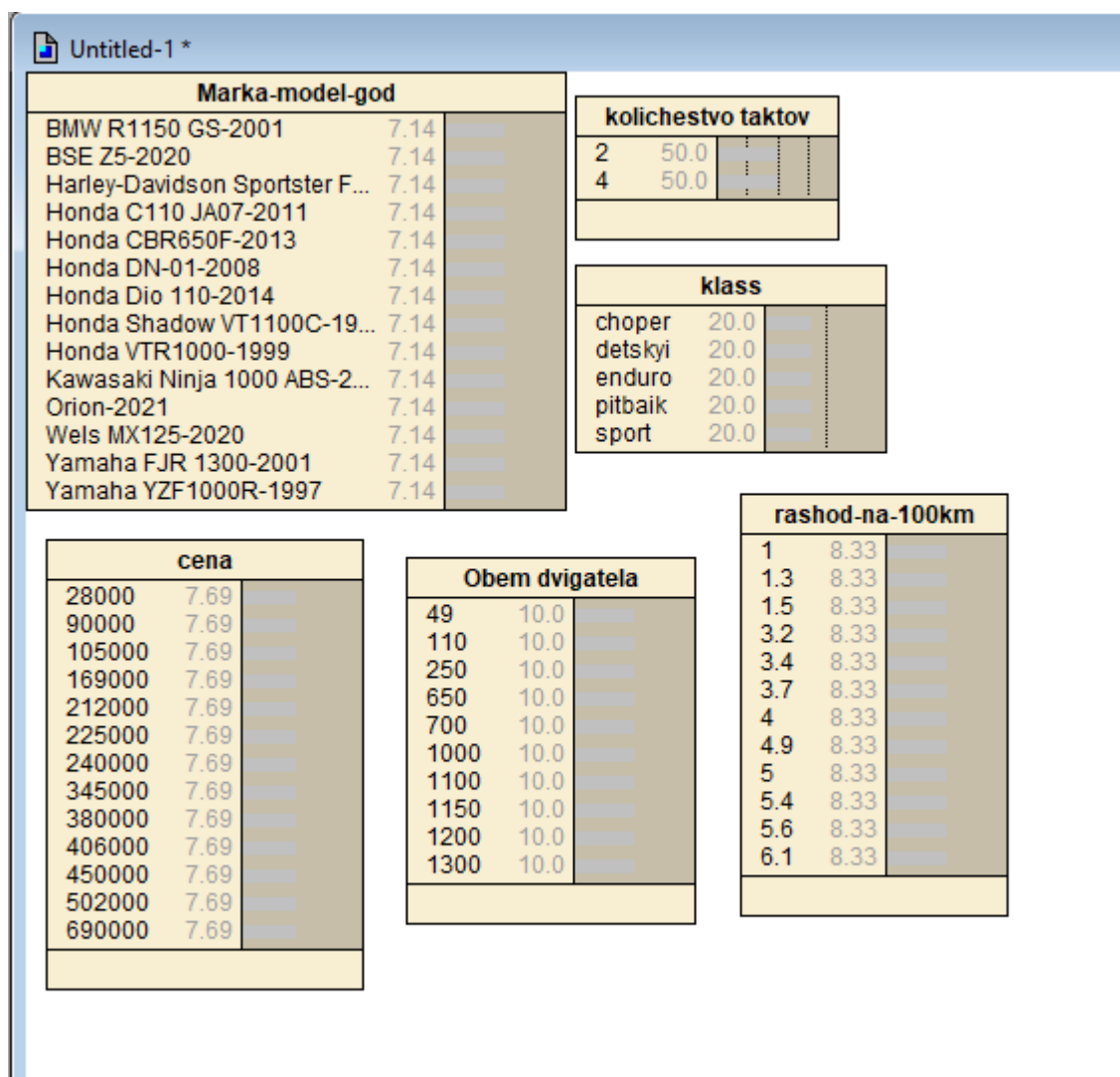


Рис. 4. Результат обучения сети

Чтобы определить систематизацию данных необходимо задать связи между узлами. Для этого выполним операцию Tree Augmented Naive Bayes. Но перед этим необходимо выбрать целевой узел. Мы выбрали klass. Так как нужно определить в первую очередь для каких целей будет использоваться мотоцикл. После чего выполнили команду Cases - Learn - Learn TAN Structure, в диалоговом окне была выбрана наша таблица Excel. Результат можно увидеть на Рис.5.

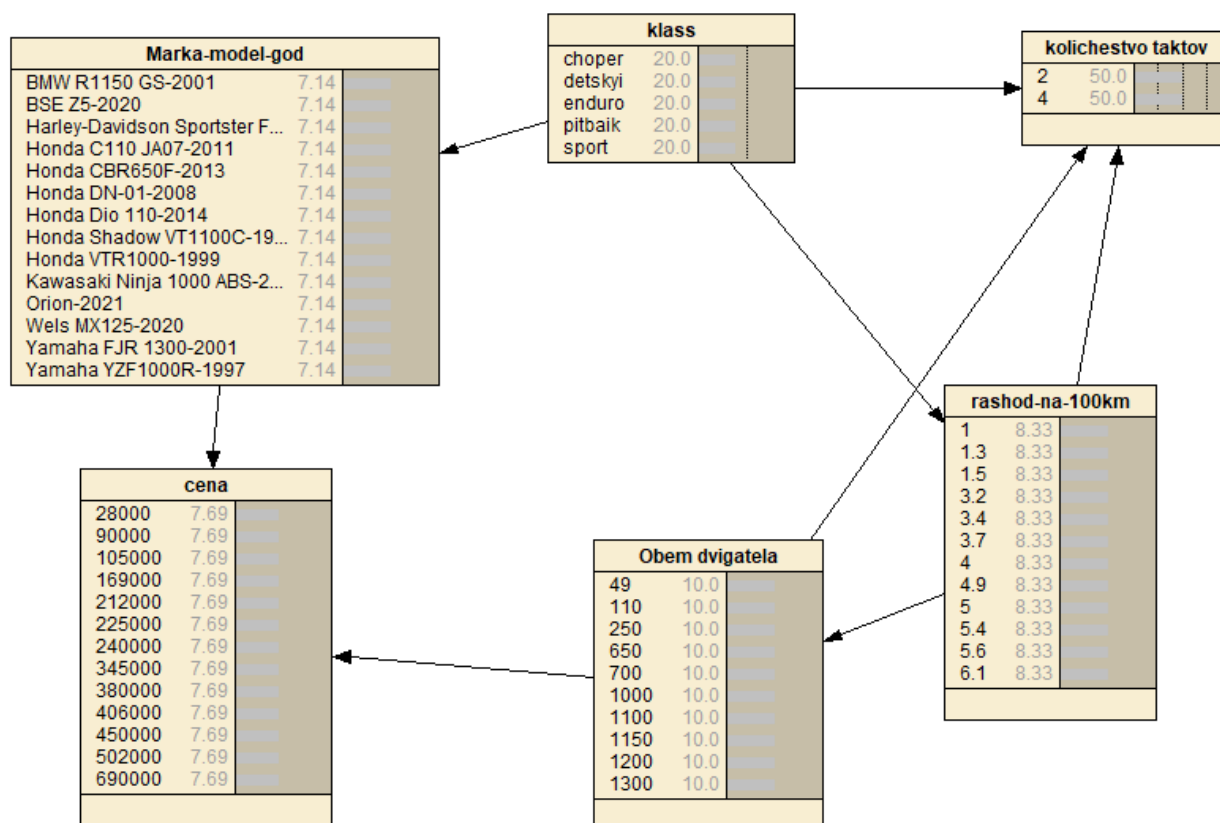


Рис.5 Добавление связей с помощью Tree Augmented Naive Bayes

Рассматривая построенные связи можно заметить, что они не совсем точны и не совсем корректно описывают нашу систему. Например, не совсем корректно будет утверждение, что количество тактов зависит от объема двигателя. Определим связи с точки зрения логических рассуждений, удалив некорректные и добавив верные суждения. Результат можно увидеть на рис.6.

Далее для получения результатов необходимо выполнить обучение полученной сети, используя EM-алгоритм. Воспользуемся клавишей в панели быстрого доступа Cases - Learn - Learn Using EM, выбираем файл MS Excel и жмем "открыть". Результат на рис. 7.

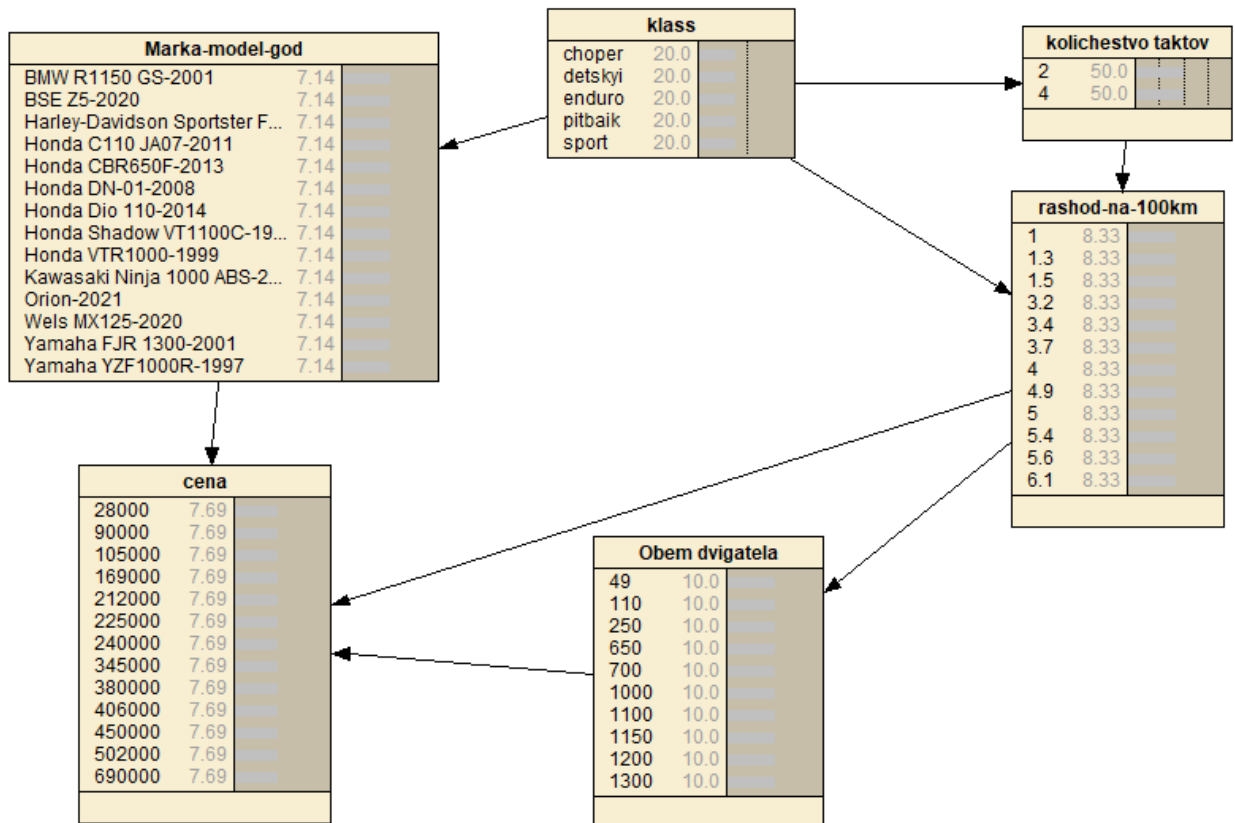


Рис. 6. Удаление противоречивых связей и оптимальный порядок узлов

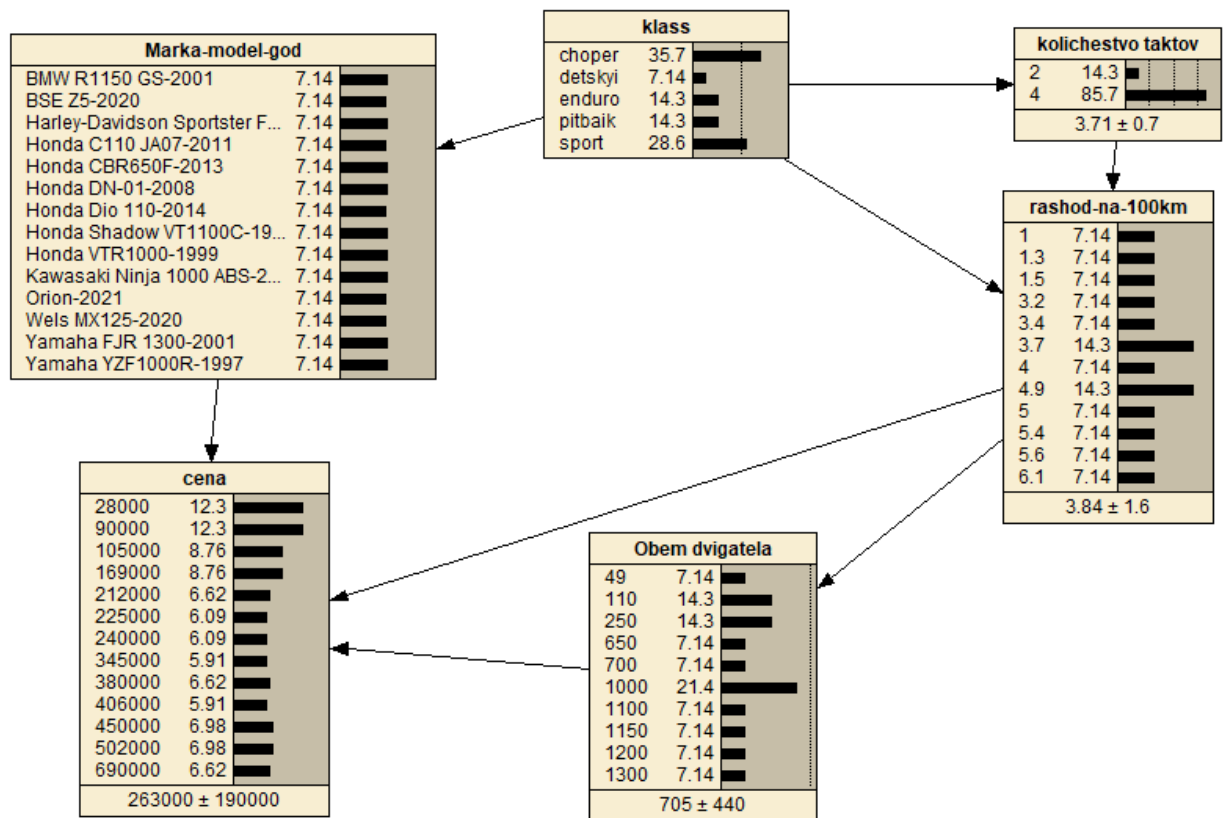


Рис. 7. Результат EM алгоритма

Применим полученную сеть для поставленной задачи. Рассмотрим различные варианты поведения сети, в процессе изменения различных факторов.

Так при выборе количества тактов 4 и расходу топлива в 1.3л/100км, мы получаем результат с несколькими моделями мотоциклов.

Результат на рис. 8.

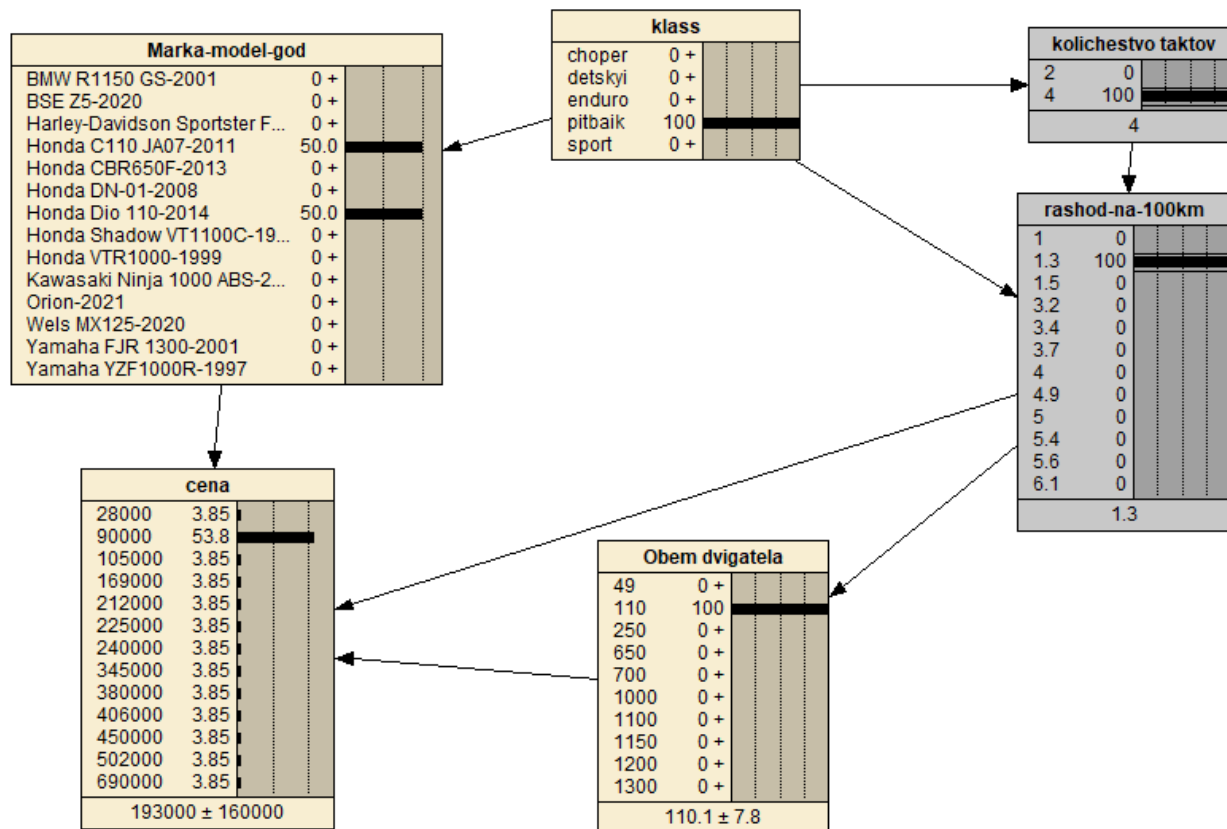


Рис. 8. Результат выбора по предпочтениям

Добавим предпочтения по таким параметрам как: цена. Таким образом, система показывает, что с вероятностью 92,2% клиенту следует выбрать следующую модель мотоцикла: Honda Dio 110-2014г. Результат на рис. 9.

Предположим, что клиент хочет выбрать спортивный мотоцикл, но при этом максимально бюджетно. Для этого он выберет класс спорт и выберет самую низкую цену. И система выдаст ему результат. По данным параметрам клиенту подходит мотоцикл: Yamaha YZF1000R-1997г. Результат выборки можно увидеть на рис. 10.

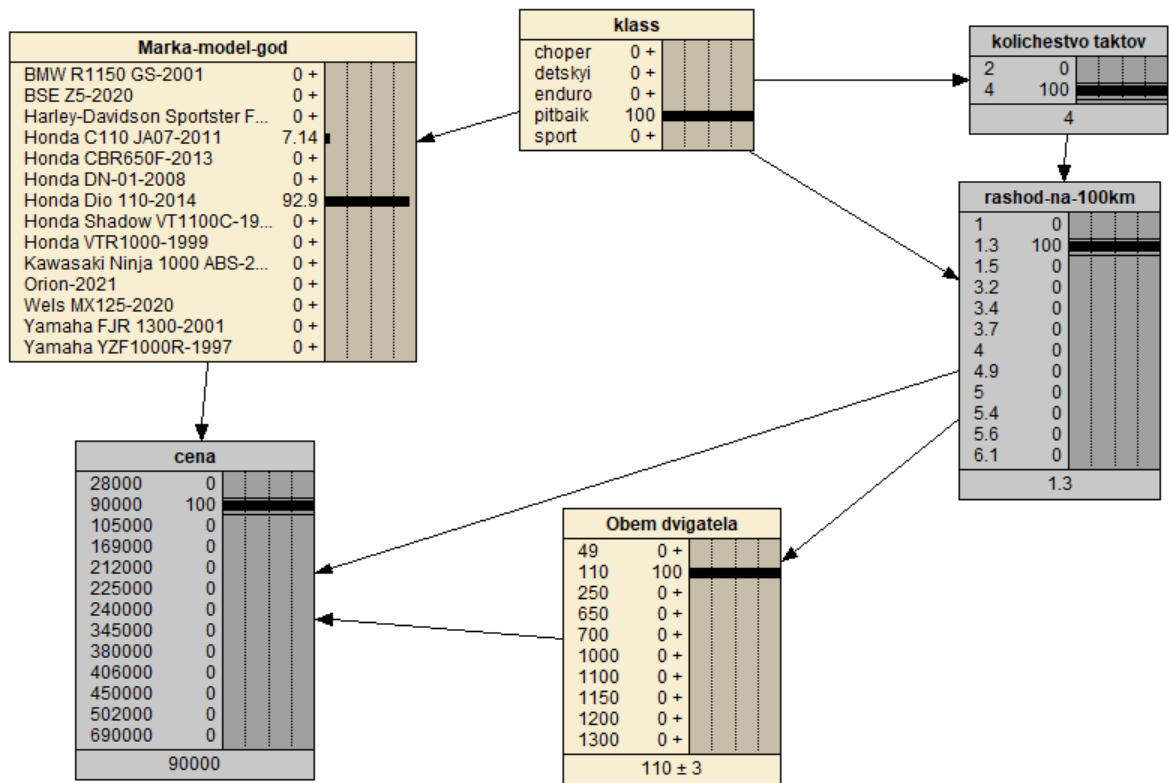


Рис. 9. Предпочтения по количеству тактов, расходу топлива и цене

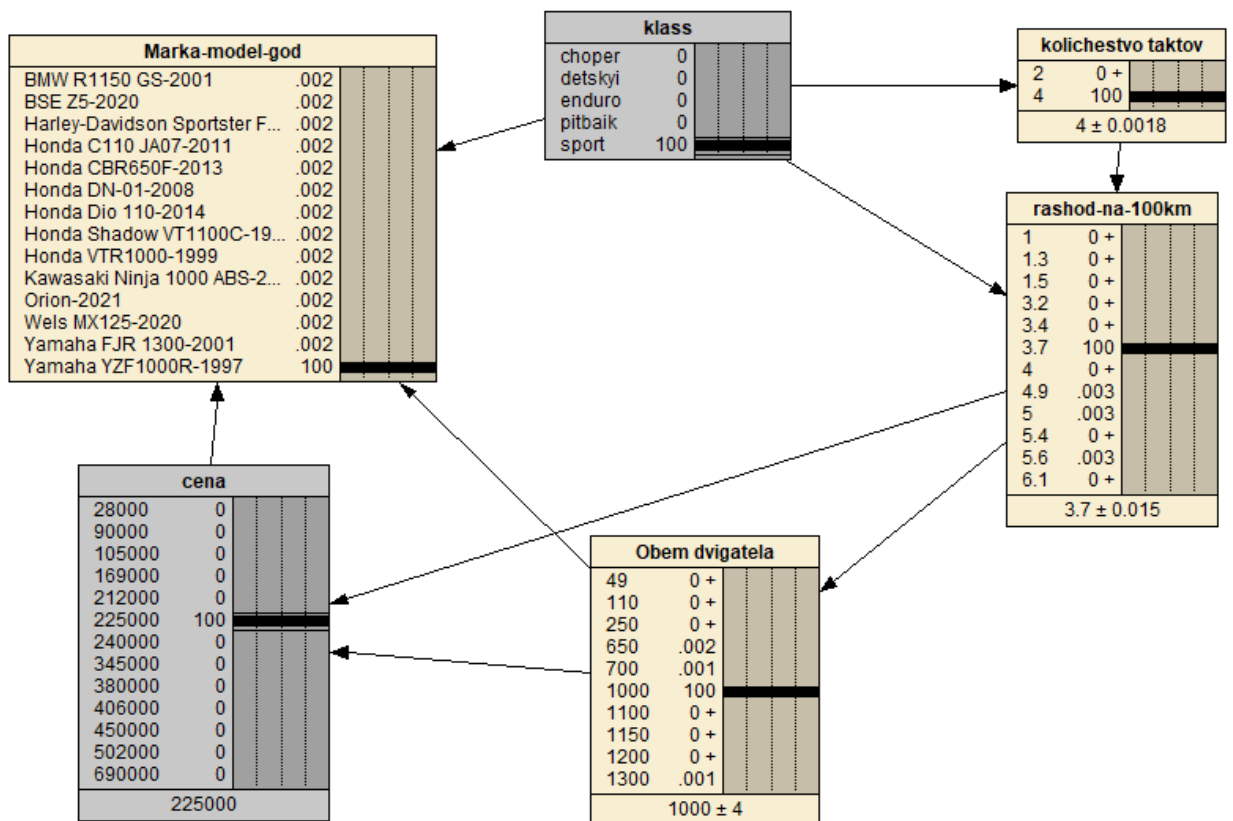


Рис.10 Выбор предпочтений клиента классу и цене мотоцикла.

Данная система позволяет определить вероятность покупки определенной модели мотоцикла из существующего списка по определенным параметрам для клиента.

Таким образом, в ходе нашего исследования была разработана байесовская система доверия, которая помогает упростить выбор покупателям в подборе комфортного мотоцикла, путем отбора характеристик по запросу клиента.

Поставленная цель была достигнута.

Библиографический список

1. Мальчевская Е.А., Бирилло А.И., Харитонов Н.А. Развитие матрично-векторного подхода в алгоритмах локального априорного вывода в алгебраических байесовских сетях // Труды СПИИРАН . 2017. С. 139-145
2. Звягин Л.С. Применение байесовского подхода в измерениях аналитических данных как фактор формирования процессов системного экономического развития // Молодой ученый. 2017. С. 256-260 URL: <https://moluch.ru/archive/156/44114/>
3. Сироткин А.В. Локальный априорный вывод в алгебраических байесовских сетях: комплекс основных алгоритмов // Труды СПИИРАН. Вып.5. СПб.: Наука, 2007. С. 100-111
4. Маклакова Г.Г. Система оценки качества услуг телекоммуникационной сети дистанционного обучения на основе байесовских сетей доверия // Новые компьютерные технологии. 2008. Т. 6. № 1(6). С. 72-73.
5. Яковлев Д.А., Синева И.С. Детектирование веб-роботов с помощью байесовских сетей. // Технологии информационного общества. 2017. С. 292-293.
6. Размахнина А.Н., Баженов Р.И. Исследование процесса принятия решений с помощью байесовских сетей доверия в системе Hugin // Постулат. 2017. № 8. С. 1-8.
7. Терентьев А.Н., Бидюк П.И., Коршевнюк Л.А. Алгоритм вероятностного вывода в байесовских сетях // Системні дослідження та інформаційні технології. 2009. № 2. С. 107-111.
8. Gan H., Zhang Y., Song Q. Bayesian belief network for positive unlabeled learning with uncertainty // Pattern Recognition Letters. 2017. Т. 90. С. 28-35
9. Phan T. D., Smart J. C.R., Capon S. J., Hadwen W. L., O. Sahin Applications of Bayesian belief networks in water resource management: A systematic review // Environmental Modelling & Software. 2016. Т. 85. С. 98-111.
10. <https://www.mdpi.com/2073-4441/11/12/2642/htm>
11. Баркалов С.А., Азаранова Т.В., Полухин П.В. Управление процессом тестирования Веб-приложений методом фаззинга на основе динамических байесовских сетей // Вестник Южно-Уральского государственного университета. 2017. С. 7-9