

ОРГАНІЗАЦІЯ МЕДИКО-ПРОФІЛАКТИЧНОЇ СЛУЖБИ

УДК 616.23+616.21: 233.4.56

А. М. Игнатъев, Т. А. Ермоленко, В. Е. Кузьмин*, А. В. Кузьмина, Т. А. Сидельникова, Л. А. Бацуля

АНАЛИЗ ВЛИЯНИЯ РАЗЛИЧНЫХ ФАКТОРОВ НА ДИНАМИКУ ИЗМЕНЕНИЯ ПЛОТНОСТИ КОСТНОЙ ТКАНИ В ХОДЕ ЛЕЧЕНИЯ ПАТОЛОГИЙ ОПОРНО-ДВИГАТЕЛЬНОГО АППАРАТА С ИСПОЛЬЗОВАНИЕМ ДЕРЕВА РЕШЕНИЙ.

Одесский государственный медицинский университет,
Физико-химический институт им. А.В.Богатского* АН Украины, г. Одесса.

В целях совершенствования диагностики патологических состояний опорно-двигательного аппарата на основе тонких изменений плотности костной ткани обработка данных денситометрии и антропологических измерений (роста и веса пациентов) проводилась методом построения дерева решений.

Стремительное развитие информационных технологий, прогресс в методах сбора, хранения и обработки данных позволил собирать огромные массивы данных, которые необходимо анализировать. Деревья решений являются методом автоматического анализа данных, способом представления правил в гираросеческой последовательной структуре, где каждому объекту соответствует единственный путь, дающий решение [1]. Под правилом понимается логическая конструкция, представленная в виде «если, то...»

Область применения деревьев решений достаточно широка. Например, в медицине – для диагностики различных заболеваний, в биомедицинских исследованиях - для определения групп, для которых лучше всего подходит данное лечение; в молекулярной биологии - для анализа строения аминокислот и т.д.

Все задачи, решаемые этим методом, могут быть объединены в следующие три класса:

- Описание данных:

Деревья решений позволяют хранить информацию о данных в компактной форме, вместо них хранят деревья решений, содержащие точное описание объектов.

- Классификация:

Деревья решений относят объекты к одному из заранее известных классов, причем целевая переменная должна иметь дискретные значения.

- Регрессия:

если целевая переменная имеет непрерывные значения, деревья решений позволяют установить зависимость целевой переменной от независимых (входных) переменных. Например, к этому классу относятся задачи численного прогнозирования (предсказание значений целевой переменной).

Идея построения дерева решений впервые была высказана Хантом в 1966 году и развита Р. Куинленом [2]

Пусть нам задано некоторое обучающее множество T , содержащее объекты (примеры), каждый из которых характеризуется t атрибутами (признаками), один из которых указывает на принадлежность объекта к определенному классу.

В таблице 1 представлены основные понятия из теории деревьев решений, которые будут употребляться в данной работе:

Таблица 1

Основные понятия теории деревьев решений

Название	Описание
Объект	Пример, шаблон наблюдений
Атрибут	Признак, независимая переменная, свойство
Метка класса	Зависимая переменная, целевая переменная,
Узел	Признак определяющего класс объекта
Лист	Внутренний узел дерева, узел проверки
	Конечный узел дерева, узел решения
Проверка (test)	Условия в узле

Пусть через C_1, C_2, \dots, C_k обозначены классы, а $1, 2, \dots, k$ – метки класса. В данном случае имеют место 3 ситуации:

-множество T содержит один или более примеров, относящихся к одному классу C_k ;

В этом случае дерево решений для T -это лист, определяющий класс C ;

-множество T не содержит ни одного примера (пустое множество). Тогда дерево решений - снова лист, но класс, ассоциированный с листом, выбирается из другого множества, отличного от T , например, из множества ассоциированного родителем;

-множество T содержит примеры, относящиеся к разным классам. В этом случае множество T разбивается на некоторые подмножества, для которых выбирается один из признаков, имеющий два или более отличных друг от друга значений O_1, O_2, \dots, O_n . Множество T разбивается на подмножества T_1, T_2, \dots, T_n , где каждое T_i содержит все примеры имеющие значение O_i ; для выбраковки признака. Эта процедура рекурсивно будет продолжаться до тех пор, пока конечное множество не будет состоять из примеров относящихся к одному и тому же классу.

Вышеописанная процедура лежит в основе многих современных алгоритмов построения деревьев решений и носит название «разделение и захват» (divide and conquer). При использовании данного метода построения дерева решений происходит сверху вниз. Поскольку все объекты заранее относятся к известным классам, процесс построения дерева решений называется обучением.

На сегодняшний день существует значительное число алгоритмов, реализующих деревья решений CART, C4.5, NewId, CHAID, CN2 и т.д.

При построении деревьев решений особое внимание уделяется следующим вопросам:

- выбору критерия атрибута, по которому пойдет разбиение;
- определению этапа остановки обучения;
- отсечение ветвей.

Общее правило разбиения заключается в следующем – выбранный атрибут должен разбить множество, так чтобы получаемые в результате подмножества состояли из объектов, принадлежащих к одному классу, или максимально приближенных к этому, т.е. количество объектов из других классов в каждом из этих множеств было как можно меньше. Для этой цели наиболее часто используются теоретико-информационный и статистический критерии.

Использование статистических методов для оценки целесообразности дальнейшего разбиения, так называемая «ранняя остановка» процесса построения дерева решений, привлекательно в плане экономии времени обучения. Однако, этот подход приводит к менее точным классификационным моделям. Поэтому вместо остановки более целесообразно пользоваться отсечением.

Очень часто алгоритмы построения деревьев решений дают сложные деревья с множеством узлов и ветвей, что затрудняет их интерпретацию.

Гораздо предпочтительней иметь дерево, состоящее из меньшего количества узлов, которым бы соответствовало большое количество объектов обучающей выборки. Для решения вышеописанной проблемы часто применяется так называемое отсечение ветвей. В отличие от процесса построения, отсечение ветвей происходит снизу вверх, двигаясь с листьев дерева, отличая узлы как листья, ибо заменяя их поддержку. Иногда даже усеченные деревья могут быть все еще сложны для восприятия. В этом случае пользуются методикой извлечения правил из дерева с последующим созданием наборов правил, описывающих классы.

Для извлечения необходимо исследовать все пути от корня до каждого дерева.

Каждый такой путь даст правило, где условиями будут являться проверки узлов, встретившихся на пути.

К преимуществам использования деревьев решений следует отнести:

- быстрый процесс обучения;
- интуитивно понятная классификационная модель;
- высокая точность прогноза, сопоставимая с другими методами (например, классификации, прогнозирования и моделирования с помощью нейронных сетей);
- построение непараметрических моделей.

Таким образом, методология деревьев решений является важным инструментом в работе специалистов, занимающихся анализом данных, и может способствовать решению вопросов практической медицины и медико-биологических научных исследований.

Материал и методы: В целях совершенствования диагностики патологических состояний опорно-двигательного аппарата на основе тонких изменений плотности костной ткани обработка данных денситометрии и антропологических измерений (роста и веса пациентов) проводилась методом построения дерева решений. Для построения дерева использовался алгоритм C8RT, являющийся алгоритмом построения бинарного дерева решений – дихотомической классификационной модели. Каждый узел дерева при разбиении имел только 2х потомков.

Полученное дихотомическое дерево включало 141 узел, 71 из, которых-концевые. Общее число уровней классификации равно 18, классификация начинается при движении от корневой вершины в зависимости от значений скорости прохождения ультразвука через костную ткань (СПУ), роста и веса пациентов. Полученное дерево решений, представленное на фиг.1 является графической классификационной схемой для диагностики патологических состояний опорно-двигательного аппарата, от начальных до тяжелых степеней патологии.

Матрица ошибок (Табл.2) показывает, что средняя ошибка разработанного дерева решений составляет 2,2%, что свидетельствует о высокой степени вероятности прогноза. Исследования были проведены на 181 пациенте в возрасте от 20 до 65 лет. Для определения плотности костной ткани был использован аппарат Achilles express американской фирмы Lunar. Измерение проводились на пяточной кости, наиболее чувствительной к метаболическим изменениям. Данные по скорости прохождения ультразвука через костную ткань (СПУ), и антропометрии приведены в таблице 3.

Таблица 3

Матрица ошибок

	Наблюдаемые категории						Всего
	0	1	2	3	4		
Предсказанные категории	0	30	0	0	0	0	30
	1	0	15	0	0	0	15
	2	0	0	5	0	0	5
	3	0	0	0	12	0	12
	4	0	0	0	0	34	34
	Всего	30	15	5	12	34	96

Количество ошибок - 0

Таблица 2.

Данные денситометрических и антропологических измерений.

№ п/п	УПП	Рост	Вес	Скорость ультразвув волны
1	0	173	108	1553
2	0	160	72	1556
3	0	167	89	1557
4	0	158	49	1555
5	0	160	89	1572
.....
177	3	173	83	1559
178	3	168	60	1523
179	3	185	76	1542
180	3	165	62	1537
181	3	165	116	1540

Степень тяжести патологии опорно-двигательного аппарата и данные денситометрических и антропологических измерений отражает модель (рис.1).

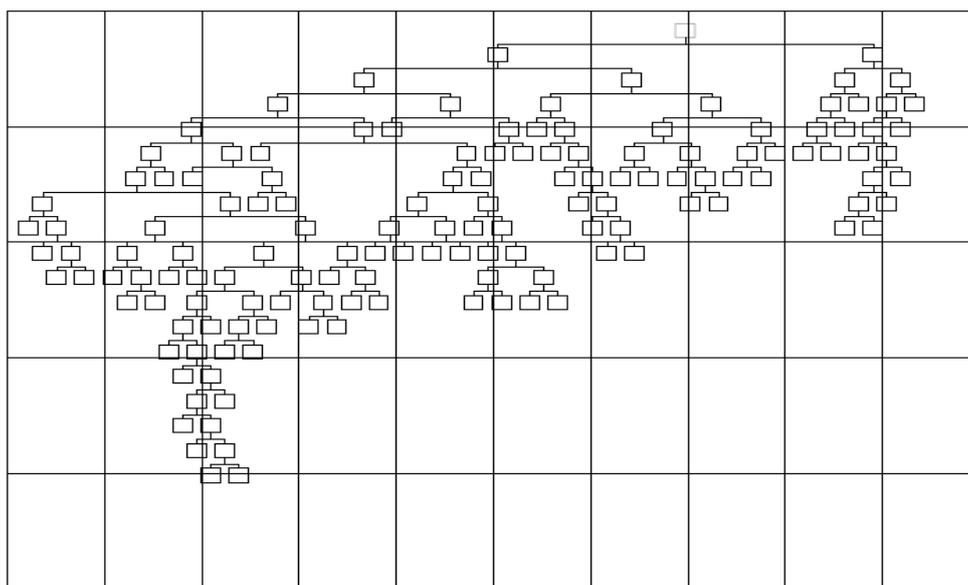


Рис.1. Схема модели диагностики патологических состояний опорно-двигательного аппарата от начальных до тяжелых степеней патологии.

Далее метод деревьев классификации был применен для оценки влияния различных факторов (возрастных, антропометрических и физиологических) на динамику изменения плотности костной ткани в ходе лечения патологий опорно-двигательного аппарата в репродуктивном и постменопаузальном периодах.

Для построения дерева решений наилучшие результаты дало использование алгоритма C8RT.

Целевыми параметрами являлись:

- исходный индекс плотности костной ткани (Stiffindex);
- изменение Stiffindex через год без медикаментозного вмешательства;
- изменение Stiffindex через полгода лечения;
- изменение Stiffindex через год лечения;

Изменения индекса плотности кости исследования по сравнению с исходным оценивались в баллах. Для составления прогноза изменения качества костной ткани учитывались следующие данные:

- возраст пациента;
- рост;
- вес;
- наличие профессиональной вредности в баллах (от 0 до 3);
- наличие/отсутствие переломов;
- количество родов;
- продолжительность кормления в баллах (от 0 до 3);
- нарушение менструации;
- болевой синдром в баллах (от 0 до 3);
- астения в баллах (от 0 до 3);

Было обследовано 96 женщин в возрасте от 15 до 48 лет.

- общая слабость в баллах (от 0 до 3);
- миома в баллах (0/1);
- урологические проблемы (0/1);
- исходный Stiffindex

Были построены следующие зависимости:

- зависимость изменения Stiffindex в течении года работы, основанная на различных возрастных, антропометрических и физиологических параметрах без медикаментозного лечения по сравнению с исходным Stiffindex.

- зависимость применения Stiffindex в течении 6 месяцев при медикаментозном лечении по сравнению с исходным.

- зависимость изменения Stiffindex в течении 1 года при медикаментозном лечении.

В последних 2-х случаях анализ проводился с учетом и без учета характера изменения индекса плотности костной ткани в первый год работы без медикаментозного вмешательства.

Полученные результаты: Полученные модели деревьев решений обеспечивают безошибочную классификацию.

Модель отражает прогноз падения Stiffindex через год без медикаментозного лечения и профилактики. Наиболее существенными факторами в этом случае являются возраст пациента, его профессиональный стаж и количество переломов. В равной степени важны при определении этой тенденции величина исходного Stiffindex, количество родов, продолжительность кормления и вес пациента.

Модель II отражает прогноз увеличения Stiffindex после лечения в течении 6 месяцев с учетом тенденции к его падению в течении года. Этот фактор, а также возраст и длительность работы во вредных условиях оказывают наибольшее влияние на понижение плотности костной ткани.

На рис. 3 представлена схема классификационного дерева, в таблице 4-матрица ошибок для моделей II и III.

На рис. 2 приведена схема классификационного дерева, а в таблице 4 – матрица ошибок.

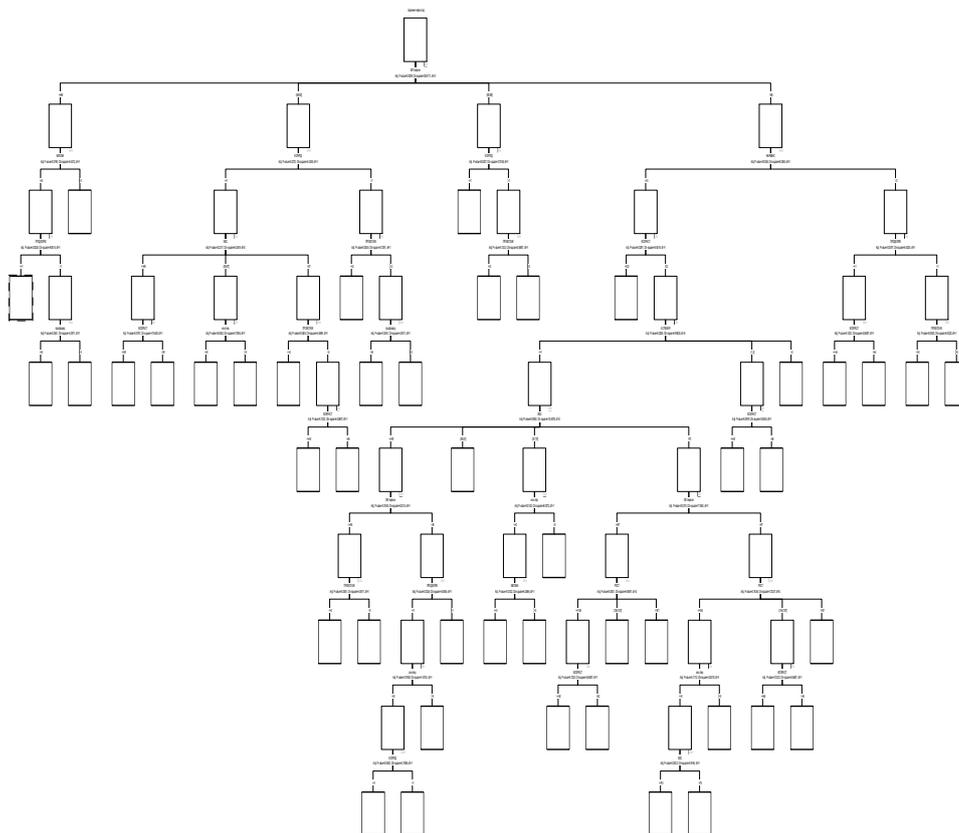


Рис. 2. Прогноз падения Stiffindex через год без медикаментозного лечения и профилактики.

Таблица 4

Матрица ошибок

		Наблюдаемые категории			Всего
		0	1	2	
Прогнозируемые категории	0	41	0	0	41
	1	0	42	0	42
	2	0	0	64	64
Всего		41	42	64	147

Количество ошибок – 0

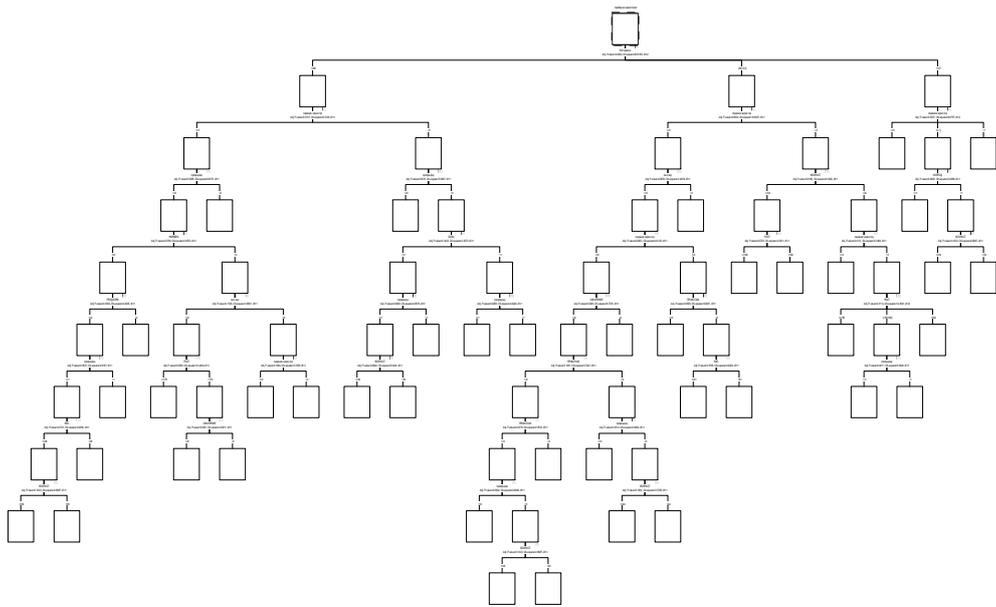


Рис.3. Прогноз увеличения Stiffindex после 6 месяцев лечения с учетом тенденции его падения в течение года.

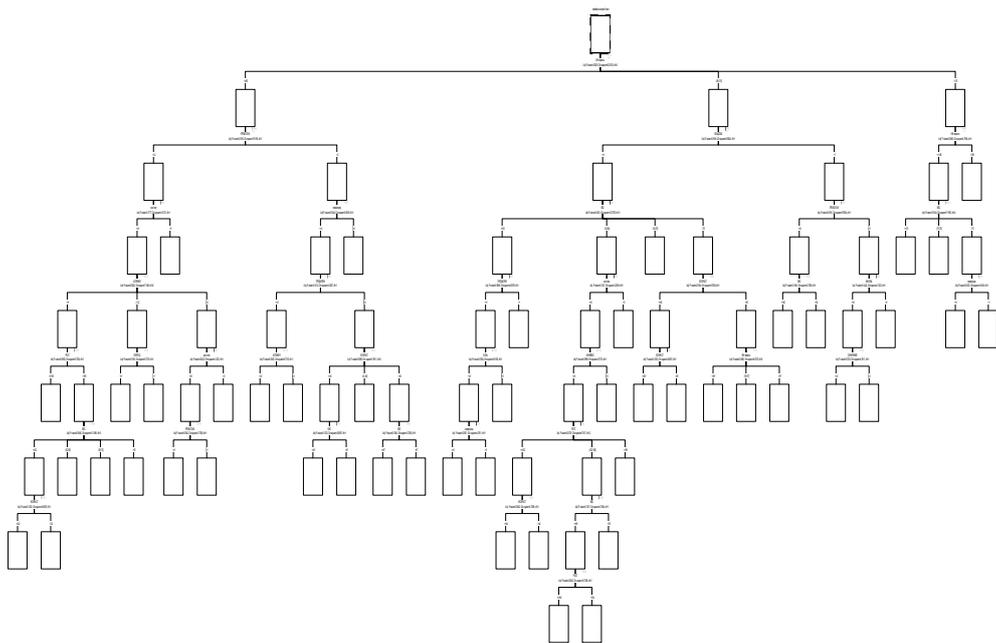


Рис. 4. Прогноз увеличения Stiffindex после 6 месяцев лечения без учета тенденции его падения в течение года.

Модель III отражает тенденцию повышения Stiffindex через 6 месяцев лечения (без учета падения этого параметра через год без лечения).

Наиболее существенными факторами в этом случае являются вес и возраст пациента, в равной степени важны первоначальный Stiffindex, рост, профстаж и вредность, определенное влияние оказывают астения, количество переломов, продолжительность кормления.

Модель IV отражает тенденцию роста Stiffindex через 12 месяцев лечения с учетом первоначальной тенденции к падению через год. В этом случае наиболее существенным фактором является первоначальный Stiffindex, затем возраст пациента и наличие профвредностей.

На тенденцию к изменению Stiffindex в равной мере влияют такие показатели, как количество родов, продолжительность кормления, количество переломов, рост, астения, а также показатель падения Stiffindex через год без лечения.

На рис. 5 представлена схема дерева классификации, а таблица 5 отражает матрицу ошибок моделей IV и V.

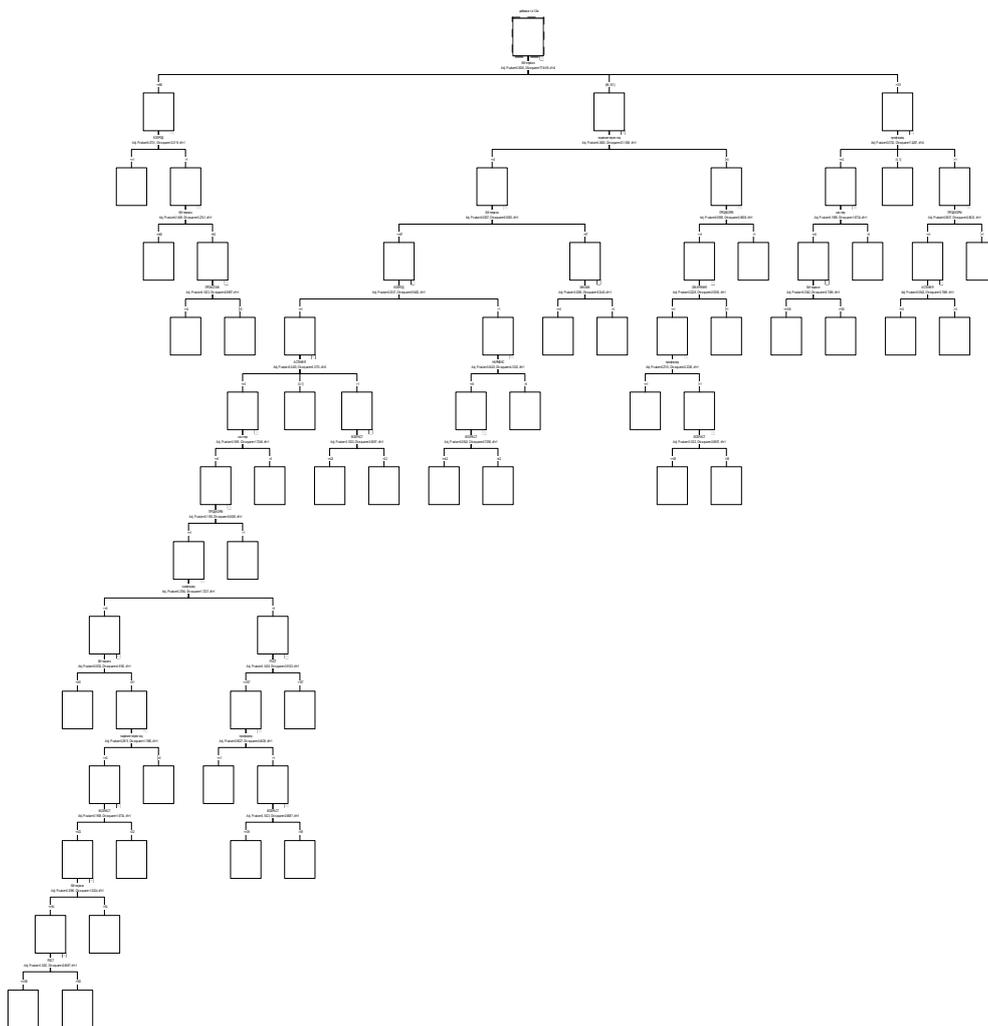


Рис.5. Прогноз увеличения Stiffindex после 12 месяцев лечения с учетом тенденции его падения в течение года.

Матрица ошибок

		Наблюдаемые категории			Всего
		0	1	2	
Прогнозируемые категории	0	41	0	0	41
	1	0	42	0	42
	2	0	0	64	64
	Всего	41	42	64	147

Количество ошибок - 0

Модель V отражает тенденцию прироста Stiffindex при медикаментозном лечении в течении 12 месяцев (без учета падения Stiffindex через год без лечения). Здесь также определяющим фактором является величина первоначального Stiffindex, затем возраст и продолжительность кормления.

Существенными показателями, влияющими на способность организма к нормализации плотности костной ткани, являются нарушение менструации и рост пациента.

На рис. 6 представлена схема дерева классификации.

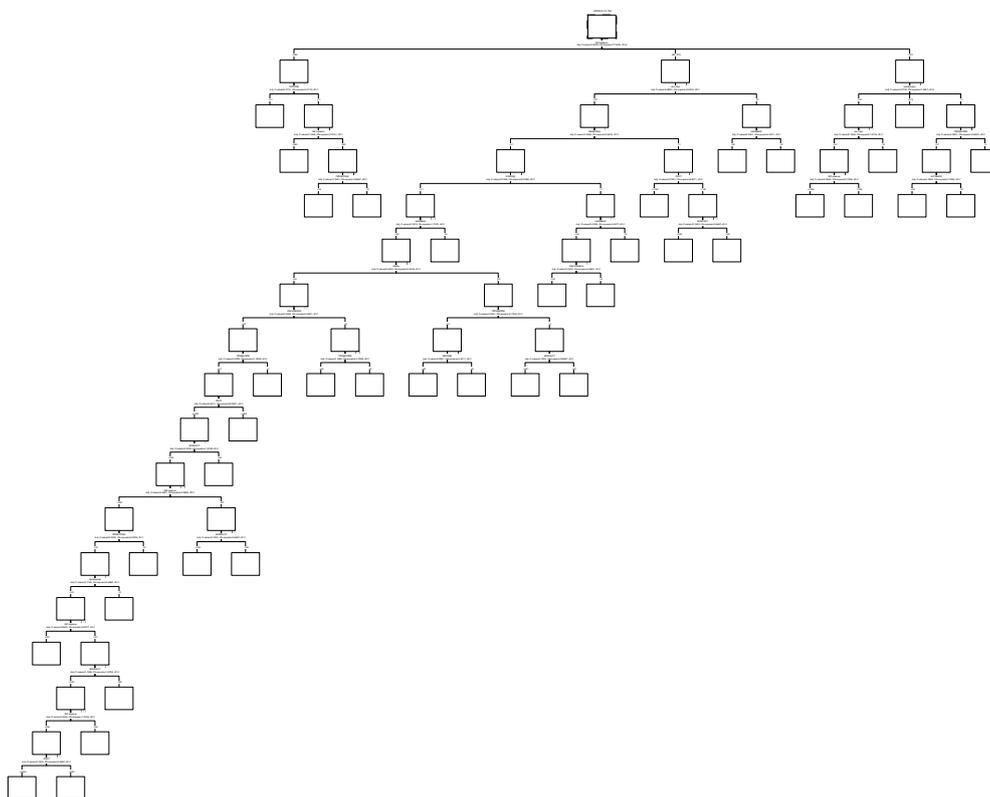


Рис.6. Прогноз увеличения Stiffindex после 12 месяцев лечения без учета тенденции его падения в течение года.

Для группы пациентов (147 женщин) была построена модель дерева классификации для диагностики патологий опорно-двигательного аппарата, содержащая 62 узла, из которых 33 концевых.

Наиболее существенными факторами являются возраст, вес и рост пациента, определяет классификацию исходной Stiffindex. Весьма существенным фактором в этом случае оказывается наличие урологических проблем, которые усугубляют проблемы нарушения фосфорно-кальциевого обмена и являются дополнительным фактором риска формирования патологии опорно-двигательного аппарата.

Вывод. Таким образом, использование метода деревьев решений дает возможность не только установить диагноз патологии с высокой степенью надежности, но и выявить факторы, оказывающие наиболее существенное влияние на состояние костной ткани пациента, факторы риска для формирования патологий опорно-двигательного аппарата.

Ключевые слова: остеопороз, факторы риска, математическое моделирование, деревья решений.

Литература:

1. Акобир Шахиди. Деревья решений – общие принципы работы /http://www.base-group.ru/trees/description.htm
2. Quinlan I. Ross: Programs for Machine learning. - Morgan Kaufmann Publ. - 1993.

Summary. Ignatiev A.M., Yermolenko T.A., Kuzmin V.Ye., Kuzmina A. B., Sidelnikova T. A.,

Batsulia L. A. Analysis of Different Factors Influence on the Dynamics of Bone Tissue Density Change in the Course of Locomotor Apparatus Pathology Treatment with the Use of the Tree of Aims

With the aim to improve diagnosis of locomotor apparatus pathology on the basis of slight changes of bone tissue density and antropological examinations (height and weight) the authors have designed the tree of aims.

УДК 621.36

Г. Ф. Кривда, А. М. Игнатъев, Р. Г. Кривда

ПРАВОВАЯ ОЦЕНКА ПРИ ПРОФЕССИОНАЛЬНЫХ НАРУШЕНИЯХ ВРАЧЕЙ

Одесский государственный медицинский университет

Согласно данным, приведенным в последнем отчете ВОЗ, Украина занимает 79-е место по уровню организации системы здравоохранения. Лучшая в мире общественная система здравоохранения создана во Франции, на втором месте – Италия, на третьем – карликовое европейское государство Сан-Марино. Кроме них в первой десятке – государство-карлик Андорра, Мальта, Сингапур, Испания, Оман, Австрия и Япония.

Из постсоветских стран: Казахстан занял 64 место, Белоруссия – 72, Литва – 73, Эстония – 77, Молдавия – 101, Армения – 104, Азербайджан – 109, Грузия – 114, Узбекистан – 117, Россия – 130, Киргизия – 151, Туркмения – 153.

Для сравнения: Китай находится на 144 месте, Индия – на 112, Польша – на 50, Албания – на 55, сообщает агентство «Washington ProFile».