

(Институт проблем информатики и управления,

Международный университет информационных технологий, г. Алматы)

МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ В ЗАДАЧАХ РАСПОЗНАВАНИЯ ПОРОД НА УРАНОВЫХ МЕСТОРОЖДЕНИЯХ

Аннотация

Рассмотрены и проанализированы методы машинного обучения с использованием технологии искусственных нейронных сетей в задачах распознавания пород на урановых месторождениях.

Ключевые слова: распознавание образов, машинное обучение, искусственные нейронные сети, искусственный интеллект, индукционный каротаж, эвристические алгоритмы, пространство признаков, петрография.

Кілт сөздер: бейне тану, машиналық оқыту, жасанды нейрондық желілер, жасанды зерде, индукциялық каротаж, эвристикалық алгоритмдер, белгілемелер кеңістігі, петрография.

Keywords: pattern recognition, machine leaning, artificial neural networks, machine intelligence, induction caritas, heuristic algorithms, space of signs, petrography.

Подготовка и интерпретация данных в задачах распознавания литологических типов является одной из важнейших в процессе добычи урана на месторождения Казахстана. По своей сути данная задача является слабоформализуемой, решение которой можно искать в рамках научного направления, именуемого «машинное обучение» [1]. Существует множество задач и успешных приложений машинного обучения.

К методам машинного обучения относится широкий класс алгоритмов, начиная от деревьев принятия решений, генетических алгоритмов, байесовских сетей и заканчивая искусственными нейронными сетями

Искусственные нейронные сети (ИНС) широко используются в задачах классификации и распознавания образов [2, 3]. Отдельный класс сетей (сети Хопфилда, Хемминга и Кохонена [4]) используется как средство ассоциативной памяти. Со времени своего возникновения в конце 50-х годов прошлого столетия (персептрон Розенблата) ИНС прочно ассоциируются с искусственным интеллектом, поскольку имитируют важные особенности естественного интеллекта – способность к обучению и ассоциативность. Обширная библиография, посвященная нейронным сетям и их приложениям, отражает неослабевающий интерес исследователей к данному направлению (например, библиография в классических изданиях [2, 3], включает более 2000 источников).

В процессе изучения аппарата нейронных сетей возникло несколько основных направлений.

Первое направление связано с возможностью применения формируемых нейронных сетей для решения классических задач вычислительной математики. Возможность такого

подхода объясняется особенностями искусственного нейрона, который представляет собой элементарный процессор, а нейронная сеть из искусственных нейронов – параллельную структуру. Возможность распараллеливания вычислений и потенциальная высокая устойчивость сети к ошибкам открыла путь к созданию аппаратно формируемых нейронных сетей.

Второе направление связано с поисками способов обучения нейронной сети и переходом от однослойной сети нейронов к многослойной. Теоретическая ограниченность решения сложных задач классификации объектов с применением однослойных сетей обоснована в книге [5]. В то же время многослойные сети не имеют таких ограничений и могут моделировать разделяющие функции практически любой степени сложности. В рамках этого направления исследуются архитектуры нейронных сетей и их способность к решению практических задач. При этом применяются многочисленные эмуляторы, например, Neuro Office, NeuroPro, Matlab, NeuroStock, Deductor, Alyuda NeuroIntelligence и другие, моделирующие работу нейронной сети на стандартном компьютере.

Третье направление связано с решением вопросов построения универсального нейронного компьютера, способного обучаться и решать после обучения сложные задачи классификации.

Массовое внимание получило второе из упоминавшихся направлений, так как оно позволяет эмулировать нейронную сеть на стандартном персональном компьютере и использовать возможности обучения нейронных сетей. Важной задачей при использовании современных ИНС с прямым распространением сигналов является формирование обучающей выборки. Обучающую выборку и набор параметров для обучения ИНС формирует исследователь. При этом в некоторых случаях исследователь руководствуется мнением экспертов. Тем самым нейронная сеть становится способной «работать как эксперт» без необходимости выявления или моделирования причинно-следственной связи факт-вывод.

Начиная с 70-х годов прошлого столетия, искусственные нейронные сети стали применяться в задачах петрографии как средство анализа каротажных данных, в литологии, оценке минерально-сырьевой базы и т.п. [6-19]. Применение нейронных сетей к решению практических задач интерпретации каротажных данных в области нефтедобычи посвящена работа [13]. В работах [15, 16] описаны некоторые результаты применения нейронных сетей прямого распространения для интерпретации данных геофизического исследования скважин при добыче урана.

Добыча урана на месторождениях Казахстана ведется методом подземного скважинного выщелачивания, который относится к числу малозатратных, экологически безвредных способов добычи [16] (рисунок 1).

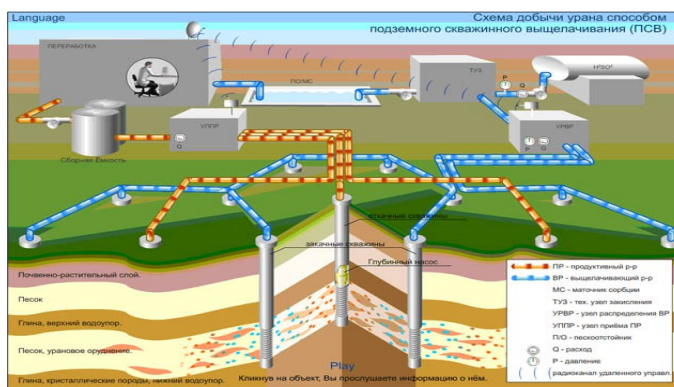


Рисунок 1 – Метод подземного скважинного выщелачивания

При этом экономические показатели процесса добычи зависят от скорости и точности интерпретации геофизических данных. В основном применяются электрические методы:

индукционный каротаж (ИК), каротаж методом кажущихся сопротивлений (КС) и методом естественной поля-ризации (ПС). Широкое применение кернового апробирования при анализе пород невозможно в силу медленности процесса получения данных. Результаты каротажа представляются в виде каротажных диаграмм, на основании которых эксперт делает заключение о глубине залегания и качестве пород (рисунок 2).

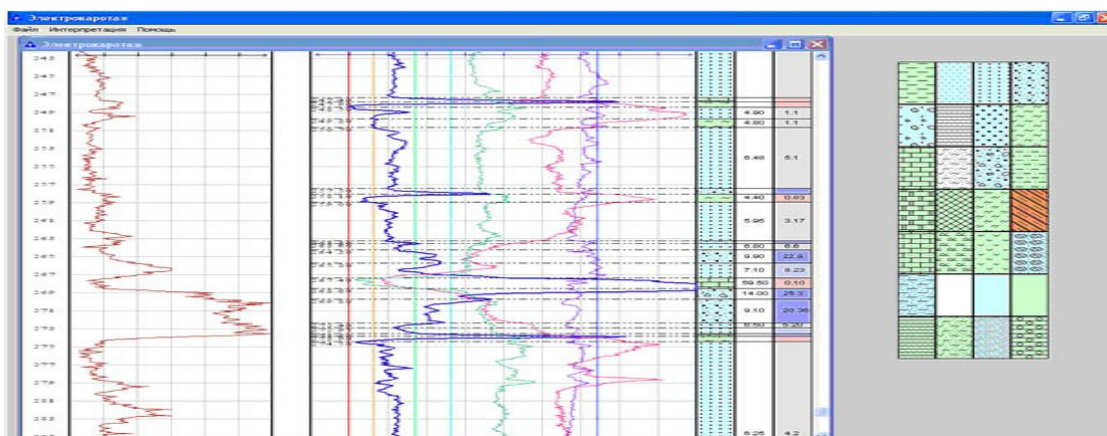


Рисунок 2 – Каротажная диаграмма

Ошибочный или неточный анализ геофизических данных приводит к потерям скважин, неоправданным трудозатратам и, в конечном счете, снижает экономические показатели добычи. Поскольку интерпретация данных каротажа носит во многом эмпирический характер, точные закономерности отсутствуют, становится очевидным возможность применения обучаемых систем, в частности, нейронных сетей. Однако, несмотря на достоинства ИНС, главная из которых это способность ИНС решать слабо-формализованные задачи, в процессе их использования имеются существенные проблемы:

- неоднозначность мнений экспертов;
- необходимость большого и равного количества примеров из разных классов;
- невозможность нейронной сети объяснить полученный результат;
- необходимость тщательной предварительной подготовки данных (очистка от аномальных значений, нормирование, сглаживание).

Перечисленные проблемы проявились на первом этапе исследования по созданию нейросетевой системы интерпретации данных каротажа на пластово-инфильтрационных месторождениях Казахстана [16, 18].

В процессе исследования выполнялась настройка нейронной сети на решаемую задачу, и анализировалось влияние следующих факторов: алгоритма обучения нейронной сети; архитектуры нейронной сети; вида нормировки; методов сглаживания; специальные методы формирования обучающей выборки; исключения сдвига каротажных данных относительно друг друга.

В общей сложности было проведено около 2 тысяч вычислительных экспериментов. В процессе экспериментов на вход сети подавалось «плавающее окно данных» с параметрами 5+1+5 (по пять точек выше и ниже текущей, плюс текущая точка) трех видов каротажа.

Разработана методика проведения экспериментов, которая включила следующие этапы:

- 1) этапы предобработки данных: выбор скважины; удаление аномальных значений; нелинейная или линейная нормировка; вейвлет анализ; устранение сдвига каротажных данных; формати-рование данных в виде «плавающего окна»;

2) этапы формирования обучающей выборки и обучения сети: объединение данных скважин или использование специального набора данных («идеальной скважины»); подбор архитектуры сети; выбор алгоритма и параметров обучения.

Были подобраны алгоритмы обучения нейронной сети, оценено влияние архитектуры, методов нормировки, сглаживания данных. Рисунки 3 и 4 иллюстрируют результаты нормировки данных.

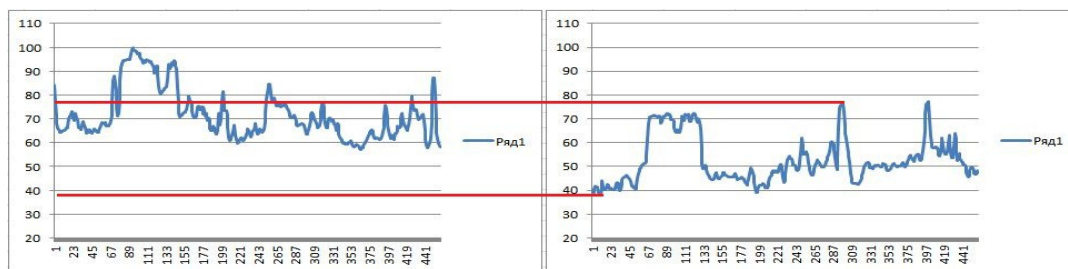


Рисунок 3 – Графики КС по двум разным скважинам

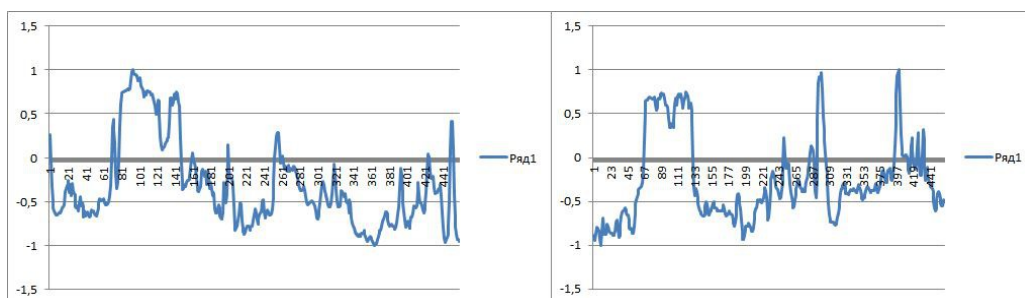


Рисунок 4 – Графики КС по двум разным скважинам после нормировки

Рисунок 5 иллюстрирует результаты сглаживания с применением вейвлета Дебоши с параметрами 1–3 (глубина разложения-масштаб).

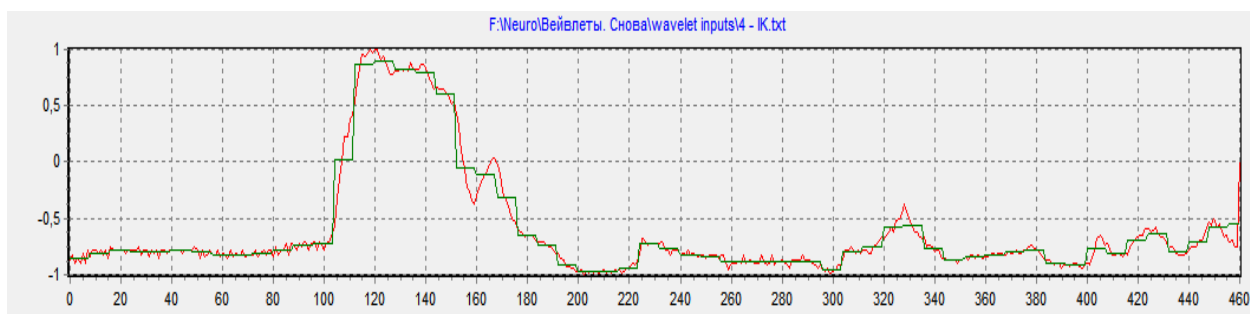


Рисунок 5 – Сигнал, сглаженный параметрами 1-3.

Красный сигнал – необработанный, зелёный – сглаженный параметрами 1-3

Анализировалось влияние сдвига каротажных данных относительно друг друга по глубине, и было разработано специальное программное обеспечение для исключения сдвига (синхронизации кривых по глубине).

В целом, результаты предварительного исследования показали, что применение исключительно ИНС на имеющемся в распоряжении исследователей массиве данных (около 12 000 отсчетов) позволяет достичь в среднем 55% степени совпадения интерпретируемых данных по сравнению с экспертами. При этом сохраняется существенный разброс качественных показателей (на некоторых скважинах процент правильных ответов составляет более 70%, на других – менее 40%). Для повышения качества интерпретации необходимо применение качественных обучающих выборок и дополнительных методов классификации. Исходный набор данных сложен для распознавания и классификации вследствие того, что объекты разных классов существенно перемешаны в пространстве признаков, а количество объектов различных классов отличается в десятки раз (рисунок 6).

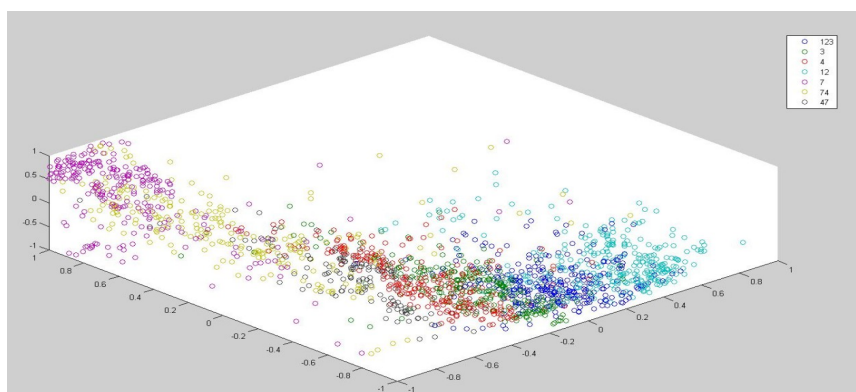


Рисунок 6 – Результаты экспертов в трехмерном (ИК, КС и ПС) пространстве признаков

Заключение. В настоящее время имеются значительные наработки в области классификации и кластеризации данных различными методами. Разработаны и применяются метрические, логи-ческие (в том числе нечеткие), байесовские, статистические, адаптивно резонансные алгоритмы (ART1,2) кластеризации и т.п. Данные методы и алгоритмы математически обоснованы. Их применение позволяет в общем случае улучшить качество обучающего множества, повысить качество распознавания данных. Результаты предварительных исследований в этой области приведены в работе [18]. Однако в результате вычислительных экспериментов показано, что применение отдельных алгоритмов, в частности, искусственных нейронных сетей, не позволяет достичь требуемой точности интерпретации данных каротажа. Применение других алгоритмов (Linear Discriminant Analysis Classifier (LDAC), Support Vector Classification (SVM) (Linear SVM и Non-linear SVM), Diagonal Linear Discriminant Analysis (DLDA), k-Nearest-Neighbor (k-NN)) обеспечивает в среднем несколько меньшую точность распознавания по сравнению с искусственными нейронными сетями.

Следовательно, необходимо создать систему, обеспечивающую качественное обучение классификатора, содержащего различные алгоритмы классификации. Иными словами, используя совокупность распознающих эвристических, в общем смысле, алгоритмов необходимо распознавать имеющиеся объекты, образующие пересекающиеся множества в пространстве признаков. Общий подход к решению таких задач изложен в [19]. Целью дальнейших исследований является создание системы интегрирующей алгоритмы распознавания различных классов и обеспечивающей повышение качественных

ЛИТЕРАТУРА

- 1 Machine learning from Wikipedia.<http://en.Wikipedia.org/wiki/>
- 2 Russell S., Norvig P. Artificial Intelligence: A modern approach. Pearson Edition, Inc., Upper Saddle River, New Jersey 07458. 2010.
- 3 Jones T. Artificial Intelligence: A Systems Approach. Infinity Science Press LLC Hingham, Massachusetts, New Delhi, 2008. ISBN: 978-0-9778582-3-1
- 4 Кохонен Т. Ассоциативная память. – М.: Мир, 1980. – 240 с.
- 5 Minsky M.L., Papert S.A. Perceptrons. An Introduction to Computational Geometry. – MIT, 1969
- 6 Baldwin J.L., Bateman R.M., Wheatley C.L. Application of a neural network to the problem of mineral identification from well logs // The Log Analyst. – 1990. – 3. – P. 279-293.
- 7 Benaouda B., Wadge G., Whitmark R.B., Rothwell R. G., MacLeod C. Inferring the lithology of borehole rocks by applying neural network classifiers to downhole logs – an example from the Ocean Drilling Program // Geophysical Journal International. – 1999. – 136. – P. 477- 491.
- 8 Saggaf M.M., Nebrija Ed. L. Estimation of missing logs by regularized neural networks // AAPG Bulletin. – 2003. – 87, № 8. – P. 1377-1389.
- 9 Yelbig K, Treitel S. Computational Neural Networks For Geophysical Data Processing / Editor: Mary M. Poulton 2001, 335 p.
- 10 Алёшин С.П., Ляхов А.Л. Нейросетевая оценка минерально-сырьевой базы региона по данным геофизического мониторинга // Новые технологии. – 2011. – № 1(31). – С. 25-32.
- 11 Raynal J.C, Serge A, Sagot A.M et al. Organization of field tests and evaluation of tricone bit performance using statistical analysis and sonic logs // Journal of Petroleum Technology. – 1971. – 23(4): 506-512.
- 12 Rogers S.J., Fang J.H., Karr C.L., Stanley D.A. Determination of lithology from well logs using a neural network. – AAPG Bulletin. – 1992. – 76(5): 731-739.
- 13 Benaouda D., Wadge G., Whitmarsh R.B., et al. Inferring the lithology of borehole rocks by applying neural network classifiers to downhole logs: An example from the Ocean Drilling Program // Geophysical Journal International. – 1999. – 136(2): 477-491.
- 14 Костиков Д.В. Инструментальные средства интерпретации геофизических исследований скважин на основе пре-образованных каротажных диаграмм с помощью многослойной нейронной сети: дис. ... к. техн. н. – М.: РГБ, 2007. – 189 p.
- 15 Kuchin Y., Muhamedyev R., Muhamedyeva L. Interpretation of log data of boreholes // The 9th International Conference: Information Technologies and Management 2011, April 14-15, 2011. – Information Systems Management Institute, Riga, Latvia, ISSN 1691-2489, 98 с.
- 16 Kuchin Y.I., Muhamedyev R.I., Muhamedyeva E.L., Gricenko P., Nurushev Zh., Yakunin K. The analysis of the data of geophysical research of boreholes by means of artificial neural networks. 5th International Conference Innovative Information Technologies for Science, Business and Education, ИТ-2012, may 10-12, 2012, Vilnius.
- 17 Яшин С.А. Подземное скважинное выщелачивание урана на месторождениях Казахстана // Горный журнал. – 2008. – № 3, ISSN 0017-2278.
- 18 Амиргалиев Е.Н., Кучин Я.В., Искаков С.Х, Мухамедиев Р.И., Мухамедиева Е.Л. Оценка качества нейросете-вого распознавания литологических слоев на урановых месторождениях // Мат-лы научно-практ. конф. «Актуальные проблемы информатики и процессов управления». – Алматы: ИПИУ. С. 262-270.

19 Журавлев Ю.И. Об алгебраических методах в задачах распознавания и классификации. Математические методы и их применение. Распознавание. Классификация. Прогноз. – М.: АН СССР, 1988. – Вып. 1. – С. 9-16.

REFERENCES

- 1 Machine learning from Wikipedia.<http://en.Wikipedia.org/wiki/>
- 2 Russell S., Norvig P. Artificial Intelligence: A modern approach. Pearson Edition, Inc., Upper Saddle River, New Jersey 07458. 2010.
- 3 Jones T. Artificial Intelligence: A Systems Approach. Infinity Science Press LLC Hingham, Massachusetts, New Delhi, 2008. ISBN: 978-0-9778582-3-1
- 4 Kokhonen T. The associative memory. – M: World, 1980. – 240 p.
- 5 Minsky M.L., Papert S.A. Perceptrons. An Introduction to Computational Geometry. – MIT, 1969
- 6 Baldwin J.L, Bateman R.M., Wheatley C.L. Application of a neural network to the problem of mineral identification from well logs // The Log Analyst. – 1990. – 3. – P. 279-293.
- 7 Benaouda B., Wadge G., Whitmark R.B., Rothwell R. G., MacLeod C. Inferring the lithology of borehole rocks by applying neural network classifiers to downhole logs – an example from the Ocean Drilling Program // Geophysical Journal International. – 1999. – 136. – P. 477-491.
- 8 Saggaf M.M., Nebrija Ed. L. Estimation of missing logs by regularized neural networks // AAPG Bulletin. – 2003. – 87, № 8. – P. 1377-1389.
- 9 Yelbig K, Treitel S. Computational Neural Networks For Geophysical Data Processing / Editor: Mary M. Poulton 2001, 335 p.
- 10 Alyoshin S. P., Poles A.L. Neyrosetevaya an assessment of mineral resources of the region according to geophysical monitoring // New technologies. – 2011. – № 1(31). – С. 25-32.
- 11 Aynal J.C, Serge A, Sagot A.M et al. Organization of field tests and evaluation of tricone bit performance using statis-tical analysis and sonic logs // Journal of Petroleum Technology. – 1971. – 23(4): 506-512.
- 12 Ogers S.J., Fang J.H., Karr C.L., Stanley D.A. Determination of lithology from well logs using a neural network // AAPG Bulletin. – 1992. – 76(5): 731-739.
- 13 Benaouda D., Wadge G., Whitmarsh R.B., et al. Inferring the lithology of borehole rocks by applying neural network clas-sifiers to downhole logs: An example from the Ocean Drilling Program // Geophysical Journal International. – 1999. – 136(2): 477-491.
- 14 Kostikov D. V. Work benches of interpretation of geophysical researches of slits on the basis of the transformed logging charts by means of a multi-layer neural network. Thesis of Candidate of Technical Sciences. – М.: RGB, 2007. – 189 p.

15 Kuchin Y., Muhamedyev R., Muhamedyeva L. Interpretation of log data of boreholes // The 9th International Conference: Information Technologies and Management 2011, April 14-15, 2011, Information Systems Management Institute, Riga, Latvia, ISSN 1691-2489, 98 c.

16 Kuchin Y.I., Muhamedyev R.I., Muhamedyeva E.L., Gricenko P., Nurushev Zh., Yakunin K. The analysis of the data of geophysical research of boreholes by means of artificial neural networks // 5th International Conference Innovative Information Technologies for Science, Business and Education, IIT-2012, may 10-12, 2012, Vilnius.

17 Jashin S.A. Underground acidic in situ leaching of uranium at the Kazakhstan deposits // "Gorny Zhurnal" Scientific-technical and industrial journal. – 2008. – N 3.

18 Amirgaliyev Ye.N. Kuchin Ya.V. Iskakov S. X, Mukhamediyev R. I. Mukhamediyeva E.L. Assessment of quality of neuronetwork recognition of lithological layers on uranium fields // Materials of the scientific and practical conference «Actual Problems of Informatics and Administrative Processes». – Almaty: IPIC. – С. 262-270.

19 Zhuravlyov Yu.I. About algebraic methods in recognition and classification tasks. Mathematical methods and their application. Recognition. Classification. Forecast. – M.: Academy of Sciences of the USSR, 1988. – V. 1. – P. 9-16.

Резюме

Е. Н. Әміргалиев, С. Х. Ысқақов, Я. В. Кучин, Р. И. Мұхамедиев

(Информатика және басқару проблемалары институты,

Халықаралық ақпараттық технологиялар университеті, Алматы қ.)

УРАН КЕН ОРЫНДАРЫНЫҢ ҮЛГІЛЕРДІ ТАҢУ ЕСЕПТЕРІНДЕ

МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІ

Уран кен орындарының үлгілерін тану есептерінде жасанды нейрондық желілер технологиясына негізделген машиналық оқыту әдістері қарастырылып талданған.

Кілт сөздер: бейне тану, машиналық оқыту, жасанды нейрондық желілер, жасанды зерде, индукциялық каротаж, эвристикалық алгоритмдер, белгілемелер кеңістігі, петрография.

Summary

Ye. N. Amirgaliyev, S. X. Iskakov, Ya. V. Kuchin, R. I. Mukhamediyev

(Institute of problems of informatics and control,
International University of Information Technologies, Almaty)

METHODS OF MACHINE LEARNING
IN TASKS OF RECOGNITION OF BREEDS ON URANIUM FIELDS

Methods of machine training with use of technology of artificial neural networks in tasks of recognition of breeds on uranium fields are considered and analyzed

Keywords: pattern recognition, machine learning, artificial neural networks, machine intelligence, induction caritas, heuristic algorithms, space of signs, petrography

Поступила 25.04.2013 г.