



Mineral Exploitation Using Neural Networks

Aysar A. Abdulrahman

University of Sulaimani, Computer Science, Kurdistan Region of Iraq
aysser.abdulrahman@univsul.edu.iq

ABSTRACT

Establishing a new site for mining operation is required to find and develop a new source of minerals with precise characteristics such as location, depth, quality and thickness. Mineral exploration is a sequential process of information gathering that assesses the mineral potential of given area. It starts with an idea of geologic model that identifies lands worthy of further exploration, and it's the one of risky and costly investments for companies. In this paper, a new method for choosing a scientific map of exploration using neural network was introduced rather than an arbitrary map. The goal of this paper is to demonstrate the effectiveness of Self Organization Map (SOM) algorithm in visual exploration of physical geographic data. The SOM is one of the most popular neural network models, which provides a data visualization technique which helps to understand high dimensional data by reducing the dimensions of data to maps. In the present paper, the algorithm is based on unsupervised learning, and Java programming codes is used to simulate the SOM algorithm.

Indexing terms/Keywords

Neural network; SOM

Academic Discipline And Sub-Disciplines

Computer Science; Neural networks

SUBJECT CLASSIFICATION

Neural Network Classification

TYPE (METHOD/APPROACH)

Experimental

INTRODUCTION

Mineral exploration is the process of finding ore (commercially viable concentrations of minerals) to mine. Mineral exploration is an intensive, organized and professional form of mineral prospecting. Mineral exploration methods vary at different stages of the process depending on size of the area being explored, as well as the density and type of information sought. Aside from extra planetary exploration, at the largest scale, is a geological mineral Province (such as the Midwest Coal fields Province of USA), which may be sub-divided into Regions. At the smaller scale are mineral Prospects, which may contain several mineral Deposits. Area selection is a crucial step in professional mineral exploration. Selection of the best and most prospective area in a mineral field will assist in making it not only possible to find ore deposits, but also to find them easily, cheaply, and quickly. Area selection is based on applying the theories behind ore genesis, the knowledge of known ore occurrences, and the method of their formation. This helps to known geological regions via the study of geological maps and to determine potential areas, where the particular class of ore deposit being sought may exist.

According to Kohonen (1990) [1], the Self-Organizing Map (SOM) is one of the most prominent models of neural network. It provides a data visualization technique which helps to understand high dimensional data by reducing the dimensions of data to a map. SOM algorithm is an unsupervised learning technique. In this paper, we examine data from coal exploration to create an effective tool for visualizing and analyzing. End product of the simulation would be a visualization map of the inputs provided to the SOM algorithm, for Illinois State.

G. Jorge (2008) [2] presented a method for clustering geo-reference data that integrates visualization of both perspectives of a SOM (output and input space) and the utilization of a three dimensional SOM for clustering geo referenced data. He finally compared the result of using the two dimensional SOM and three dimensional SOM. It is considered as a promising technique for exploratory analysis of data which is difficult to analyze because of the size of datasets, their complexity, scaling problems, and hidden patterns. The Geo-information science is used to reduce the complexity of organizing and supporting user's information exploration to find information needed more effectively and intuitively. Koua (2003) [3] has used the SOM algorithm for exploratory data analysis, information visualization, and knowledge discovery; which enables users to enhance the understanding of geospatial analysis results for fast and effective exploration of the data.

SELF-ORGANIZING MAP (SOM)

For many years, artificial neural networks (ANNs) have been studied and used to model information processing systems based on or inspired by bio-logical neural structures. Then, not only can provide solutions with improved performance when compared with traditional problem solving methods, but also give a deeper understanding of human cognitive abilities. Among various existing neural network architectures and learning algorithms, Kohonen's self-organizing map, which is one of the most popular neural network models. It is developed in 1982 by Tuevo Kohonen, a professor emeritus of the Academy of Finland. It has named "self-organized" because no supervision if required and it is learning on its through unsupervised competitive learning. Moreover, the word "Map" is because it attempts to map its weights to conform



to the given input data. The SOM can provide topologically preserved mapping from input to output spaces. Although the computational form of the SOM is very simple, numerous researchers have already examined the algorithm and many of its problems, nevertheless research in this area goes deeper and deeper there are still many aspects to be exploited [4]

Self-Organizing Map Algorithm

The SOM is an algorithm used to visualize and interpret large high-dimensional data sets. Typical applications are visualization of process states or financial results by representing the central dependencies within the data on the map. The map consists of a regular grid of processing units, "neurons". A model of some multidimensional observation, eventually a vector consisting of features, is associated with each unit. The map attempts to represent all the available observations with optimal accuracy using a restricted set of models. At the same time the models become ordered on the grid so that similar models are close to each other and dissimilar models far from each other.

Fitting of the model vectors is usually carried out by a sequential regression process, where $t = 1, 2, \dots$ is the step index: For each sample $\mathbf{x}(t)$, first the winner index c (best match) is identified by the condition

$$\forall i: \|\mathbf{x}(t) - \mathbf{w}^c(t)\| \leq \|\mathbf{x}(t) - \mathbf{w}^i(t)\|.$$

After that, all model vectors or a subset of them that belong to nodes centered around node $c = c(x)$ are updated as

$$\mathbf{m}_i(t+1) = \mathbf{m}_i(t) + h_{c(x),i}(\mathbf{x}(t) - \mathbf{m}_i(t)).$$

Here $h_{c(x),i}$ is the "neighborhood function", a decreasing function of the distance between the i th and c th nodes on the map grid. This regression is usually reiterated over the available samples [5].

METHODOLOGY

The rapidly increasing volume of digital geographic data is overwhelming for conventional analysis techniques and methods. Therefore, new approaches are needed to transform data into information, ultimately, and knowledge. In the present work, the Illinois coal geology map is used as primary data. Figure 1 shows the large scale of coal deposit in Illinois State. With the help of this figure, coal distribution geology map, Illinois State is divided in 30 square shape zones. Then, for each zone, possibilities of existing coal deposit on the base of previous study are assigned. The possibility of (1) shows the existent of the coal seam and the possibility of (0) shows that there is no coal seam. Figure 2 shows the different zone on Illinois map, and table 1 is the possibility and coordinates of different zones.

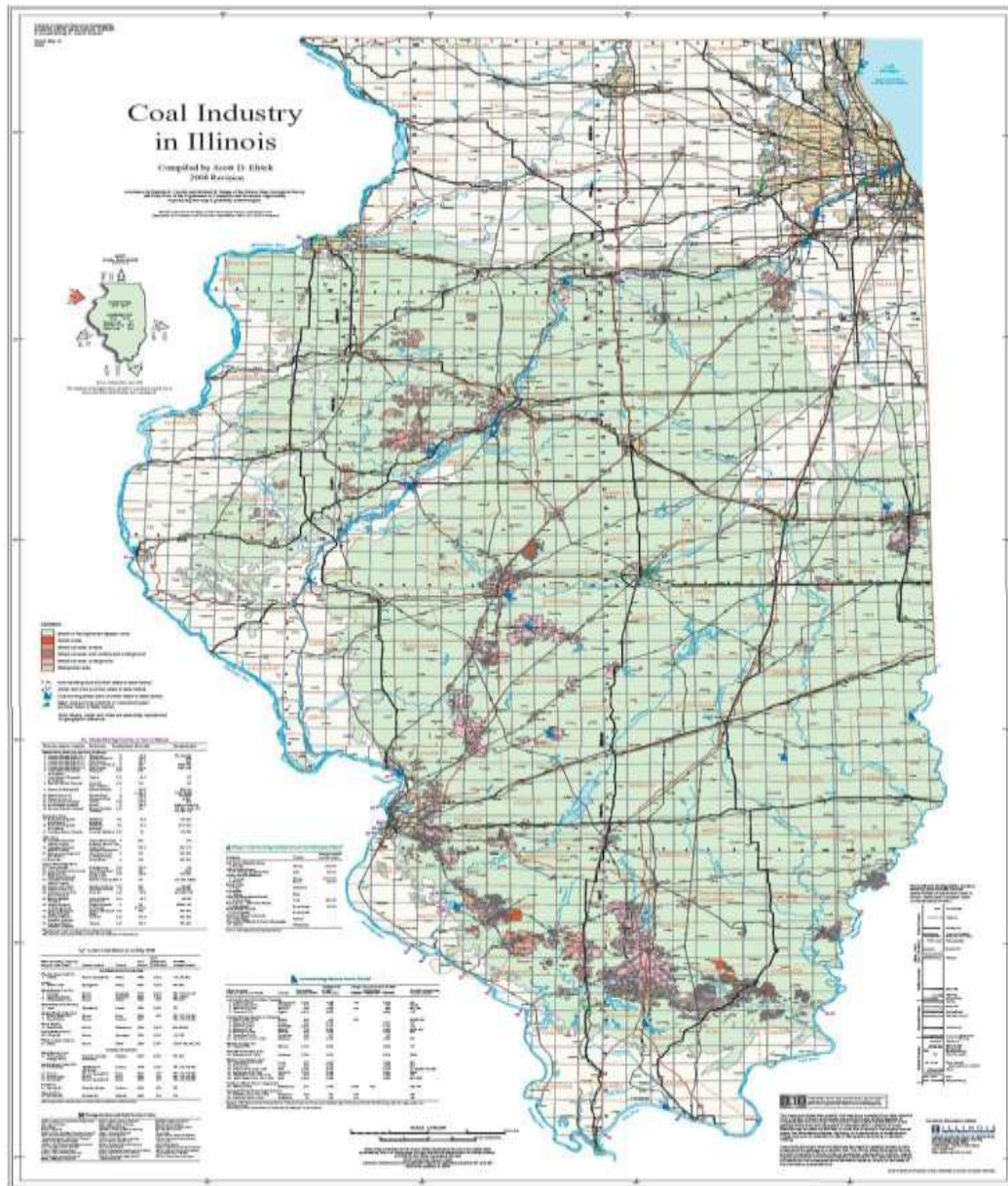


Fig 1 Illinois State Coal seam distribution (ICCI 2003)

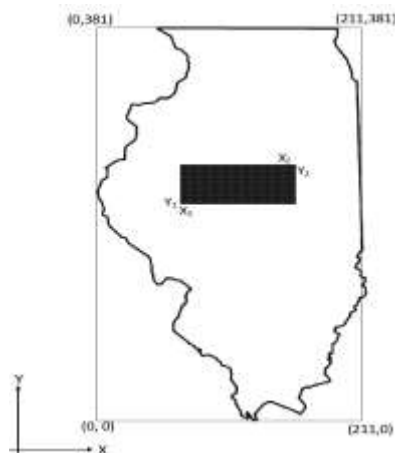


Fig 2 Illinois State model layout with dimension

Figure 2 shows the model layout with dimensions, which are used in this work.

Table 1. Possibility of exciting coal (X_1 , X_2 , Y_1 , and Y_2 are defined in figure 2)



Area#	x1	x2	y1	y2	pr
Area01	87.5	158	0	31.5	0.1
Area02	87.5	111	31.5	63	0.75
Area03	111	158	31.5	63	1
Area04	60	80	63	94.5	0.9
Area05	80	105.5	63	94.5	0.2
Area06	105.5	123.5	63	94.5	0.8
Area07	123.5	180	63	94.5	0.2
Area08	35	60	94.5	126	0.9
Area09	60	200	94.5	126	0.2
Area10	18	180	126	190	0.15
Area11	180	211	126	190	0.85
Area12	0	70	190	202.5	0.3
Area13	70	87.5	190	202.5	0.95
Area14	87.5	211	190	202.5	0.25
Area15	15	87.5	202.5	221.5	0.15
Area16	87.5	105	202.5	221.5	0.85
Area17	105	187	202.5	221.5	0.35
Area18	187.5	211	202.5	221.5	0.15
Area19	18	89	221.5	254	0.15
Area20	89	101.5	221.5	254	0.6
Area21	101.5	187	221.5	254	0.35
Area22	187	211	221.5	254	0.73
Area23	23	89	254	286	0.1
Area24	89	101.5	254	286	0.6
Area25	101.5	211	254	286	0.1
Area26	30	211	286	318.5	0.1
Area27	0	149	318.5	350	0.1
Area28	149	160	318.5	350	0.57
Area29	160	211	318.5	350	0.1
Area30	32	211	350	381	0.1

IMPLEMENTATION

Self-Organizing map algorithm is relatively difficult to understand and implement. Matlab was used to simulate the basic Kohonen algorithm just for two neurons and some random points. Although it was not straight forward to implement the algorithm in Java, but it has a better functionality for visualizing the evolution of the data grid. Therefore, the SOM algorithm is implemented using Java program. We designed two components in Java which runs simultaneously in two threads: (1) the SOM algorithm, and (2) Display Frame. SOM algorithm runs the main algorithm, while display frame access the nodes from the algorithm and create a visual two dimensional (2D) map.

Pseudo code of the algorithm is presented in Table 2. We set the initial values for learning rate and distance, and run the algorithm for (i) iterations. In each step of iteration, and for all inputs, we find the best node and update the best node and its neighbors.

Table 2. Pseudo Code in Java program

```

Algorithm Self-Organizing Map( $\alpha_0, d_0, i, inputs$ )
1:  $\alpha_0 = \text{initial learning rate}$ 
2:  $d_0 = \text{initial distance}$ 
3:  $i = \text{number of iterations}$ 
4: For step 1 to i do
5:    $d = d_0 * (1 - \text{step}/i)$ 
6:    $\alpha = \alpha_0 * (1 - \text{step}/i)$ 
7:   For all inputs (j)
8:      $bestNode = \min_k ||input - nodes_k||$ 
9:     update  $d$  neighborhood of  $bestNode$  with  $\alpha$  learning rate
10:  End For
11: End For

```

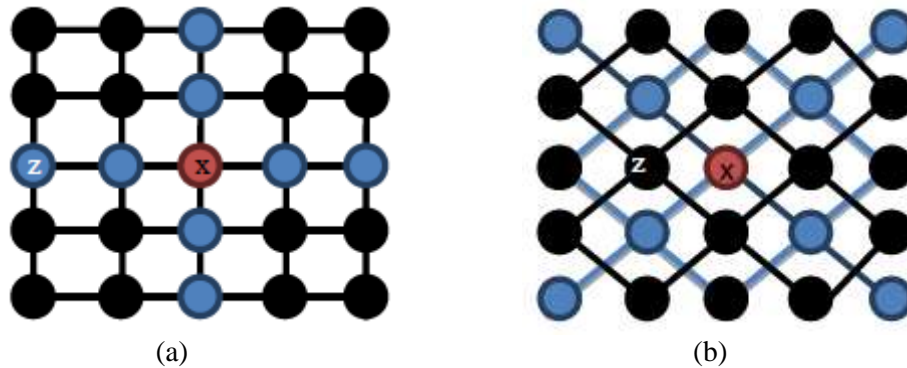


Fig 3 SOM neighborhood pattern

Figure 4 shows the sample of inputs file. The program reads the file line by line. Each parameter in a line is separated by a comma and have the following format: Area#,x1, y1, x2, y2, pr. The value of pr is the probability of finding coal in a particular area. The inputs are generated based on the probability of finding coal. A coordinate(x, y) is generated randomly, and is selected pr percent of time for a particular coordinate. Therefore, number of inputs in an area will be proportional to the probability of finding coal in that area. (Figure: 5) shows the screenshot of the program.

```

Area1,87.5,158,0,31.5,0.1
Area2,87.5,111,31.5,63,0.75
Area3,87.5,158,31.5,63,1.0

```

Fig 4 Sample of input data



Fig 5 The program Screenshot

RESULTS AND DISCUSSION

Figure 6 shows the evolution of the map for three different state of the evolution process. Initially, nodes are arranged randomly. As the learning takes place, nodes started arranging itself. We can see the final map which calculated for $i=1000$. Also notice that the learning rate and the distance for updating the node decreases as the number of iterations increases.

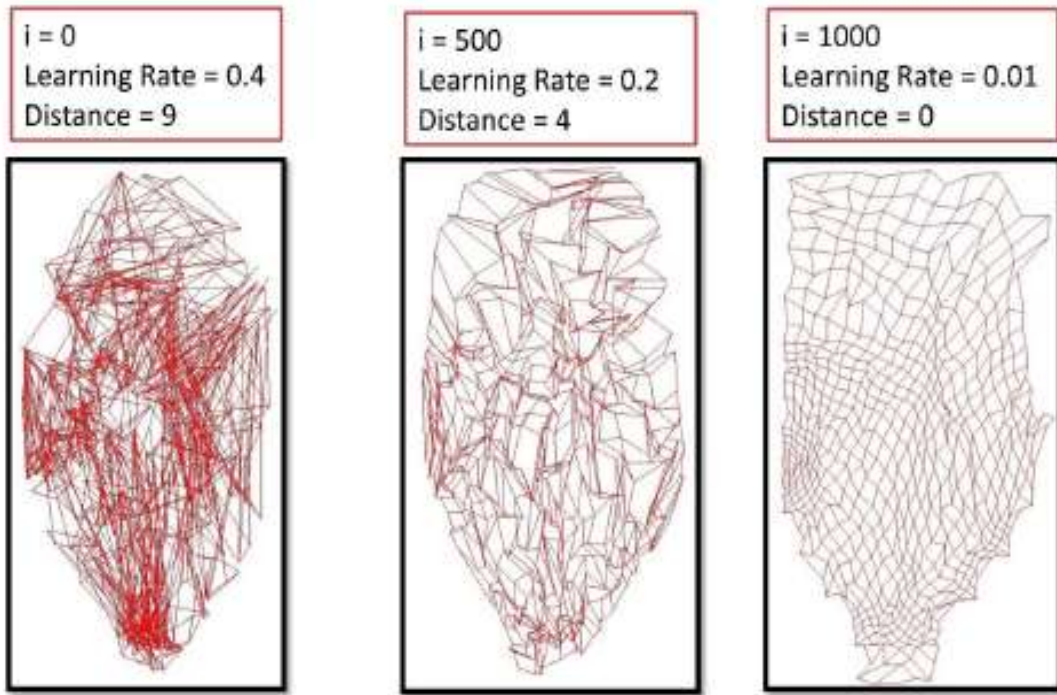


Fig 6 Evolution of the nodes

Figure 7 shows the results of the simulation, along the side of Illinois map. It is evident from the figure that the nodes are getting arranged depending on the inputs. Such arrangement is better for visualizing the data. Areas with higher coal percentage are densely populated compared to area with lower coal probability. Such visualization can be used to provide guess work for relatively unknown area based on the nodes in neighboring area, which might or might now contain coal.

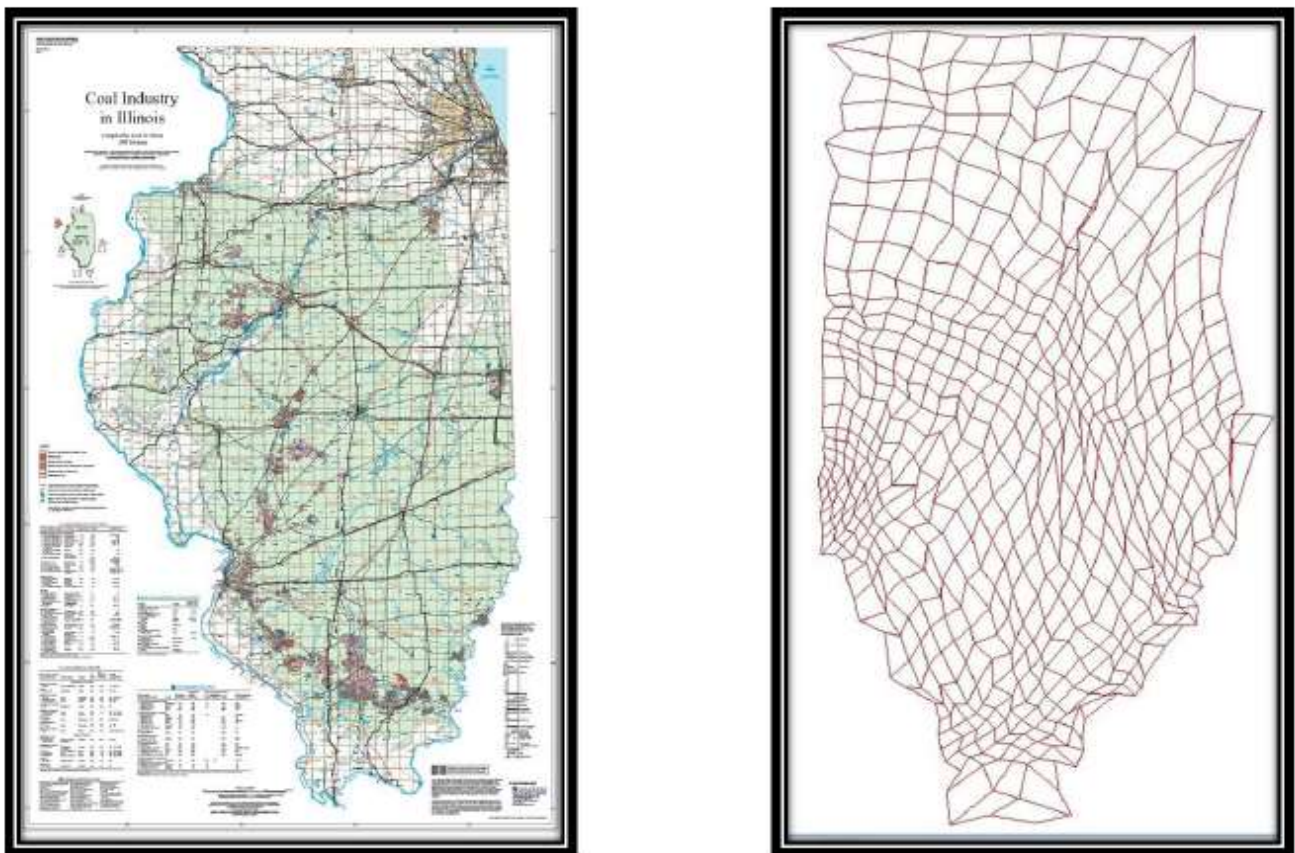


Fig 7 Final result comparisons

CONCLUSION AND FUTURE WORK

The aim of this work was to show how effective the SOM algorithm can be in exploratory data analysis of mining data and, furthermore, suggest exploration regions in a case study. The emphasis on local factors and coal distribution map are the main justifications for SOM application, in this study. New method for mapping exploration was developed. This study demonstrates the effectiveness of SOM application in visual exploration of physical geography data.

Information extraction from SOM output remains a major bottleneck. Improvements in visualization tools and interaction are needed. In this study 2D visualization was considered; and for better understanding and selecting best map, 3D visualization would be handy.

REFERENCES

- [1] T. Kohonen, "The Self-Organizing Map". Proceedings of the IEEE, vol. 78, pp. 1464-1480.(1990).
- [2] G. Jorge, "Visualization of Clusters in Geo-referenced Data Using Three-dimensional Self-Organizing Maps", University of Nova de Lisboa, (2009).
- [3] E.L. Koua, "Using Self-Organizing Maps for Information Visualization and Knowledge Discovery in Complex Geospatial Datasets", International Cartographic Conference (ICC), (2003).
- [4] T. Kohonen, "Self-organized formation of topologically correct feature map", Biological Cybernetics, Vol 43, (1982)
- [5] S. Haykin. Neural Networks and Learning Machines. 3rd ed. New York: Prentice Hall/Pearson, 2009
- [6] Obermayer, Klaus. Self-organizing Map Formation: Foundations of Neural Computation. Cambridge, Mass.: MIT Press, 2001.
- [7] Hancock, John M. "Self-Organizing Map (SOM, Kohonen Map)." Dictionary of Bioinformatics and Computational Biology, 2004.
- [8] Obermayer, Klaus. Self-organizing Map Formation: Foundations of Neural Computation. Cambridge, Mass.: MIT Press, 2001.
- [9] Villa, Alessandro E., and W. Duch. Artificial Neural Networks and Machine Learning -- ICANN 2012 22nd International Conference on Artificial Neural Networks, Lausanne, Switzerland, September 11-14, 2012, Proceedings, Part II. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.
- [10] Babloyantz, A. Self-Organization, Emerging Properties, and Learning. Boston, MA: Springer US :,1991.

Author' biography with Photo



Aysar A. Abdulrahman received his B.Sc. in Information and Communication Engineering from University of Baghdad, in 2003. In 2008, he received the M.Sc. degrees in Computer Science from University of Sulaimani. During 2008-2010, he worked as an assist lecturer at University of Sulaimani\ College of Science\ Computer Department. In Dec 2013, he got his Ph.D. in Computer Engineering from Southern Illinois University. Now, he is working as a lecturer in the University of Sulaimani\ College of Science\ Computer Department.