

Supply Chain Modelling Based on Twelve Related Features: A Novel Iteration Feature Selection Method



Hussein Alsteif, Murat Akkaya

Abstract: Real-time prediction of hour-based order entry has been lacking in literature. Compared to previous research on supply chain problems, our proposed approach overcomes the constraints of operations management with longer time periods such as weekly and monthly by developing a novel iteration model. We performed experiments on 100 products with high cumulative volume over time. Using 3 different dataset, our proposed model proved efficient in forecasting skewed demand signals with lot of noise in supply chains.

Keywords: Supply Chain, Machine learning, Noise reduction, Iteration Model, Learning algorithm.

I. INTRODUCTION

With the expansion of the activities of industrial enterprises internally and externally, many views and administrative practices have emerged that seek to describe and analyze these activities and the interactions, relationships and flows (financial, informational ...) they produce between all the parties contributing to the production process until the finished product reaches The final customer, through the various stages and paths that this product takes, so that each stage forms a link linked to the stage (the cycle) before and after it, i.e. on the basis that each product takes a specific path within these circles and in an interconnected and integrated chain before providing it (supplying it), and in This framework emerged what is called "supply chain management", and if every organization seeks to expand (or create) its market share and improve its position towards competitors and as an input to improve its production performance, it needs to evaluate the effectiveness of its supply chain management. Hence, the study seeks to A description and interpretation of the concept and stages of the supply chain in addition to providing a benchmark measure of its effectiveness based on objective criteria that are accurate and credible. The field of supply chain management is considered one of the most important modern research trends in business management because there are justifications related to economic performance that constitute an entry point for researchers, institutions and decision-makers to study the so-called "supply chain", and through this topic the study seeks to determine the true and accurate meaning of this emerging aspect.

Manuscript received on June 21, 2021.

Revised Manuscript received on September 05, 2021.

Manuscript published on September 30, 2021.

Hussein Alsteif, Girne American University, Girne, North Cyprus. E-mail: hsteif@gmail.com

Murat Akkaya, Girne American University, Girne, North Cyprus.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Management by delving into various theoretical concepts associated with it. Theoretically, this research contributes to a larger body of work that focuses specifically on operational challenges and bottlenecks in e-commerce market management. Previously, researchers have considered supply chain demand and forecasting more broadly, particularly in relation to the bullwhip effect. For example, Chan et al. (2015) created a fuzzy time series method to predict supply chain disruptions, and Bottanin et al. (2019) combined two artificial neural networks to model wholesale distribution operations.

Real-time prediction of hour-based order entry has been lacking in the literature. Compared to previous research on supply chain problems, our proposed approach overcomes the constraints of operations management with longer time periods such as weekly and monthly. In today's e-commerce market climate, from the perspective of logistics service providers, the arrival of e-commerce orders is more complex and periodic than traditional, high-volume, frequently arriving logistics orders. Due to the arrival of e-commerce orders, it is now necessary to perform demand forecasting at a much shorter interval to generate a more useful decision support tool for practitioners who make operational decisions on a regular basis.

In this paper, we present a prediction method using machine learning that incorporates three features of time series data, namely AR, MO and MA. In this way, the model can be applied to a variety of vertical industries to improve operational decision making and prediction accuracy. A four-step model implementation process and a formal model evaluation framework are provided to enable practitioners to practically apply the proposed model and systematically evaluate the predictive accuracy of their neuro-fuzzy models in their real business environment.

By introducing machine learning algorithms, this paper provides perspectives for both researchers and practitioners on how Big Data from the cloud can be better utilised to improve management decision making in B2B and B2C e-commerce market environments.

The remainder of this paper is organised as follows. Section 2 is devoted to an examination of current research and hypotheses on the topic. Section 3 presents the proposed methodology for modelling e-order arrival, accompanied by a discussion of the model implementation through a case study in Section 4. Section 5 discusses the conclusions and implications. The study's findings are presented in Section 6.



II. LITERATURE REVIEW

Many different definitions have emerged according to the different orientations and views of each researcher of supply chain management, which seeks to develop and define a comprehensive concept that focuses on the most important aspects covered by the supply chain management in the economic institution, and the following is a presentation of the content of each of them: Definition of Yves Pimor: "Supply chain management is the sequence of suppliers who contribute to the production and delivery of a good or service to the end user. This includes, by default, all the following aspects of work: Sales processing, Production, Management Inventory management: Material supply, Distribution, Procurement, Forecasting, Consumer service. According to this definition, the supply chain consists of a group of different parties and the integrated tasks necessary to provide products. (goods or services) to their end user. Alexandre Samii definition: The supply chain is: "a group consisting of three (or more) institutions connected to each other directly through the so-called upstream and downstream flows of products, services, funds and information, from the point of origin to the point of final consumption." Accordingly, supply chain management, according to Alexandre K. Samii, requires controlling the outgoing and incoming flows of the components of the supply chain with attention to the side of the feedback between these parties. Definition of "Ishaq Mahmoud Al-Shaar": The supply chain represents: "all activities related to the flow of products and services, and information from sources of supply to the final consumer." This definition focuses on the activities that constitute the supply chain and translated in the form of flows towards the end user without concern for reverse flows. (From the customer to the institution or from the institution to the suppliers). Definition of "Donald Waters": The philosophy of supply chain management is based on: "The idea of partnership in the marketing channel while ensuring a high degree of connectivity between the entities forming this channel. The idea of supply chain management is also based on maximizing profits, so that the objectives of the organization and the parties The ones that deal with them often conflict and one of them may cause harm or losses to the other party, so it is necessary to coordinate and adjust the achievable common goals in order to achieve the common goals, and here it can be noted that this definition refers to the concept of supply chain management from a marketing perspective (Marketing channel) with an indication of the importance of managing the channel in a way that achieves the problematic parties, that is, the primary goal of supply chain management is to solve the problem of conflict of objectives between the organization and the parties that deal with it. Definition of "Alan Harrison, Remko van Hoek": It is: "Planning and controlling all operations from customers to suppliers of raw materials that link partners within the framework of the supply chain in order to serve and meet the needs of the end customers." Accordingly, The goal of meeting the needs of the customer from the point of view of "Alan Harrison, Remko van Hoek" is the starting point of supply chain management and the basis of planning and control processes for the activities and operations of the

organization and is the basis of the relationship between the organization and its suppliers. Definition of Arawati Agus: Supply Chain Management includes: "Integration of the vision, culture, operations, and strategies of the organization to optimize the flows (of high value and quality) of raw materials in a reliable and innovative manner with the aim of designing and manufacturing products with quality and competitive price, and includes all activities that include providing The right product for the right consumer, in the right quantity and at the right time," illustrated through this definition the strategic dimension of supply chain management, by sharing the vision and culture and building lasting relationships with the components of the supply chain. This would create a competitive advantage as a result of controlling the prices of high quality products presented in an innovative way with Indication of the necessity to anticipate customer demands for products and provide them at the right time and place.

1. A Machine Learning Predictive Model Incorporating Time Series Data for Supply Chain.

The machine learning approach for predicting the arrival of e-orders is presented in this section with a case study for illustration. First, the technical aspect of ANFIS models is implemented, followed by the determination of the input variables, namely AR, MO and MA.

1.1. Background case

The case company has been in the e-commerce logistics industry since 2013. Although the number of partnered e-retailers has grown steadily, three e-retailers accounts for the majority of e-commerce orders. In this context, the three individual retailers are involved in the process of model building for e-order arrival prediction. Two typologies are created with a total of eight ANFIS-based prediction models to successfully evaluate the prediction output of each prediction model.

Figure 1 illustrates the background and rationale for creating two typologies. We then use this as a starting point to show how our proposed predictive approach is composed.

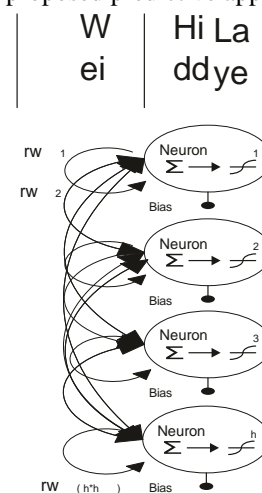


Fig 1. Background and Reasonings of Typologies



1.2. Determination

Three types of variables are proposed as determinants of the time-series E-order arrival number in this analysis. The reasons for selecting these input variables are explained below. Table 1 shows the steps we took to incorporate these variables into the creation of an ANFIS-based predictive model for predicting e-commerce order arrival.

III. DATASET

We have included data from manufacturers who use integrated ERP systems that track every product or service that enters and leaves the system. Both transactions are monetary in nature. Two manufacturers provided demand data obtained from their ERP systems. The first company makes chocolate from cocoa beans, which must be ground, converted to cocoa, and then mixed with other ingredients according to a recipe. The geographic scope of the dataset is North America. We looked at the top 100 products with the highest cumulative volume over this time, as these are the most important products for which forecast accuracy is critical. The second company specializes in toner cartridges for photocopiers and laser printers and other related products. The regional scope of the data is North America. The demand data derived from the ERP method contains 65 cycles of data. Only the top 100 items were selected for the experiment because they are the most significant.

With a total number of 3,369 active products in their scheme, the 100 products account for less than 3% of all products. However, these top brands account for 38.35% of the manufacturer's sales volume.

We used time series data on manufacturer demand collected from Statistics Canada. This is publicly available historical data covering a large portion of the manufacturing sector in the Canadian economy to add validity to the experiment.

This dataset allowed us to experiment with a wide range of demand trends across a variety of industries (214 industry categories). We used the most recent "new orders" observations, totalling 148 phases over 12 years and four months of outcomes. To maintain consistency with the other two datasets, we randomly selected 100 categories. Because of their aggregate presence and the long data collection period, trends in the data from Statistics Canada differ from those of individual manufacturers. Due to the aggregation effect, the variability is much lower than that of the individual manufacturers.

IV. EXPERIMENT

We used a representative collection of conventional prediction techniques as the control group and a set of machine learning techniques as the treatment group. Each technique from each group was used to compare the two groups. Demand was predicted one month in advance for all 100 series for the three previously defined data sets. For each technique studied, this resulted in a series of 4,700 forecast points for the chocolate manufacturer, 6,500 forecast points for the toner cartridge manufacturer, and 14,800 forecast points for the dataset Statistics Canada. However, as all forecasting methods are based on historical data, there was a fixed start-up period which slightly reduced the number of forecast observations.

In addition, the demand time series was formally divided into two sets: Preparation and Research. This is particularly true for machine learning techniques, where the training set was used to teach ML models about demand trends, and the testing set was used to determine how well the forecasting capacity generalized into the future. For each forecast data point, absolute error (AE) was used as the primary output metric. As a result, a set of absolute error values was created for a given forecasting technique. We normalized the absolute error by dividing it by the standard deviation of the training set to make it comparable for all commodities. As a result, we compared the output of different techniques using the normalized absolute error (NAE). We trained on 80% of the time series data and tested on 20% of the data. Then, using the MATLAB 7.0 setting, we used all the selected techniques to generate forecasts (MathWorks, 2005b).

V. RESULTS

The performance of the selected models on three datasets is discussed in this section. The MAE of all tested prediction techniques as applied to the chocolate manufacturer's dataset is presented in Table 1 in ascending order of error, with the best performing techniques at the top and the worst performing techniques at the bottom. The results for the toner cartridge manufacturer's dataset (Table 2) and the manufacturer's dataset Statistics Canada are presented in a similar format (Table 3). From these results, we can see that one of the methods from ML, the SVM under the supersized modelling approach, consistently outperforms the others on all three datasets. If we disregard the supersized models, the results of previous studies and the very large M3 competition are essentially replicated, suggesting that simple techniques outperform more complex and sophisticated approaches.

For example, exponential smoothing performed best on the two primary datasets of interest, chocolate, and toner cartridge manufacturer. The toner cartridge dataset was so noisy, or the patterns changed so much over time, that even exponential smoothing with a fixed parameter of 20% performed better (Table 2-Rank 3) than automatic smoothing (Table 2-Rank 5), which optimized the parameter for the training set. This basically means that automatic exponential smoothing outperforms the data, which is unexpected.

Table 1. Performance of Forecasting Techniques for Chocolate Manufacturer's Dataset

CODE	Sim1	Sim2	Sim3	Avg Val
SCM01	0.960174	0.603906	0.415931	0.660003
SCM02	0.718872	0.288407	0.608731	0.53867
SCM03	0.019592	0.610328	0.567456	0.399125
SCM04	0.109705	0.14703	0.901539	0.386091
SCM05	0.737314	0.166009	0.276117	0.393147
SCM06	0.804274	0.109855	0.479064	0.464398
SCM07	0.624367	0.358355	0.48292	0.488547



SCM08	0.755769	0.549612	0.939677	0.748353
SCM09	0.324981	0.078402	0.391003	0.264795
SCM10	0.953769	0.769319	0.488848	0.737312
SCM11	0.555659	0.208416	0.422001	0.395359
SCM12	0.191729	0.713539	0.580774	0.495347
SCM13	0.430326	0.921993	0.26751	0.539943
SCM14	0.565742	0.627821	0.628586	0.607383
SCM15	0.745956	0.743115	0.808706	0.765926
SCM16	0.938366	0.345055	0.117541	0.466987
SCM17	0.031022	0.615385	0.073614	0.240007
SCM18	0.508067	0.4338	0.080794	0.340887
SCM19	0.835502	0.054383	0.584056	0.491314
SCM20	0.617641	0.219558	0.284119	0.373773
SCM21	0.606696	0.085806	0.102048	0.26485
SCM22	0.875194	0.606128	0.559059	0.680127

Table 2. Performance of forecasting techniques for toner cartridge manufacturer’s Dataset

CODE	Sim1	Sim2	Sim3	Avg Val
SCM01	0.813264	0.019193	0.21174	0.348066
SCM02	0.032374	0.733285	0.953434	0.573031
SCM03	0.582364	0.387535	0.520776	0.496892
SCM04	0.885416	0.657341	0.730154	0.757637
SCM05	0.932262	0.28069	0.490226	0.567726
SCM06	0.652771	0.916445	0.977478	0.848898
SCM07	0.291129	0.557023	0.374393	0.407515
SCM08	0.762417	0.541032	0.329056	0.544168
SCM09	0.955649	0.223877	0.253647	0.477724
SCM10	0.315534	0.74495	0.795526	0.61867
SCM11	0.804604	0.272257	0.857037	0.644633
SCM12	0.046142	0.921378	0.480865	0.482795
SCM13	0.438076	0.778689	0.034234	0.417

Table 3. Performance of forecasting techniques for statistics Canada manufacturing Dataset

CODE	Sim1	Sim2	Sim3	Avg Val
SCM01	0.532069	0.455585	0.88218	0.623278
SCM02	0.05901	0.344746	0.206409	0.203388
SCM03	0.80302	0.763382	0.168452	0.578285
SCM04	0.761534	0.899605	0.96414	0.875093
SCM05	0.33956	0.206507	0.374575	0.30688
SCM06	0.47721	0.149166	0.669831	0.432069
SCM07	0.854196	0.364242	0.565091	0.59451
SCM08	0.040887	0.186284	0.476311	0.234494
SCM09	0.197964	0.603761	0.050311	0.284012
SCM10	0.940489	0.228951	0.291204	0.486881
SCM11	0.358935	0.753283	0.068006	0.393408
SCM12	0.376949	0.351663	0.54209	0.423568
SCM13	0.844027	0.608653	0.449925	0.634202
SCM14	0.95728	0.398858	0.943595	0.766577
SCM15	0.680664	0.048792	0.504974	0.411477

SCM16	0.873052	0.470882	0.958222	0.767386
SCM17	0.001302	0.508953	0.165677	0.225311
SCM18	0.94449	0.465009	0.888831	0.76611
SCM19	0.92387	0.087243	0.870228	0.627113
SCM20	0.617653	0.582812	0.091225	0.430563
SCM21	0.955342	0.772799	0.676136	0.801426
SCM22	0.401159	0.971921	0.645603	0.672894

The moving average method, set to a window of 6 periods (Table 2-Rank 4), yielded the same result. The automated versions had similar overfitting problems and performed worse (Table 2-Rank 7) than the manual versions (Table 2-Rank 4). For the two manufacturer datasets, the mean error of the automated exponential smoothing was 0.7446, while the mean error of the fixed 25% exponential smoothing was 0.7331. With the mean of the automatic exponential smoothing, the moving average with a window of 6 cycles had an average error of 0.7612 and a mean difference significance of 0.2116. For all three data sets, the average error of the automatic exponential smoothing was 0.6234, while the average error of the fixed exponential smoothing of 25% was 0.6527. The average error of the moving average with a six-period window was 0.6118. Thus, the automatic exponential smoothing, the 25% exponential smoothing, and the moving average with a six-period window all showed similar performance.

VI. CONCLUSION AND DISCUSSION

The aim of this research was to investigate the applicability and benefits of machine learning techniques for forecasting skewed demand signals with a lot of noise in supply chains. While practitioners have access to a variety of forecasting algorithms, there are few objective and repeatable guidelines for which method should be used. We have shown empirically in this study that the best conventional approach for a producer is automatic exponential smoothing with the series' first value as the initial value. We also discovered that many of the more advanced machine learning techniques perform poorly, likely due to the limited number of historical time periods available for each product. When learning and predicting single time series, none of the machine learning techniques can consistently outperform the best conventional equivalent (exponential smoothing).

REFERENCES

1. Bogue, R. (2016), "Growth in e-commerce boosts innovation in the warehouse robot market", *Industrial Robot: International Journal*, Vol. 43 No. 6, pp. 583-587.
2. Bottani, E., Centobelli, P., Gallo, M., Kaviani, M.A., Jain, V. and Murino, T. (2019), "Modelling wholesale distribution operations: an artificial intelligence framework", *Industrial Management and Data Systems*, Vol. 119 No. 4, pp. 698-718.
3. Boyacioglu, M.A. and Avci, D. (2010), "An adaptive network-based fuzzy inference system (ANFIS) for the prediction of stock market return: the case of the Istanbul stock exchange", *Expert Systems with Applications*, Vol. 37 No. 12, pp. 7908-7912.



4. Boysen, N., de Koster, R. and Weidinger, F. (2018), "Warehousing in the e-commerce era: a survey", *European Journal of Operational Research*, Vol. 277 No. 2, pp. 396-411.
5. Catalao, J.P.D.S., Pousinho, H.M.I. and Mendes, V.M.F. (2011), "Hybrid wavelet-PSO-ANFIS approach for short-term electricity prices forecasting", *IEEE Transactions on Power Systems*, Vol. 26 No. 1, pp. 137-144.
6. Chan, F.T., Samvedi, A. and Chung, S.H. (2015). "Fuzzy time series forecasting for supply chain disruptions", *Industrial Management and Data Systems*, Vol. 115 No. 3, pp. 419-435.
7. Chang, J.R., Wei, L.Y. and Cheng, C.H. (2011), "A hybrid ANFIS model based on AR and volatility for TAIEX forecasting", *Applied Soft Computing*, Vol. 11 No. 1, pp. 1388-1395.
8. Colla, E. and Lapoule, P. (2012) "E-commerce: exploring the critical success factors", *International Journal of Retail and Distribution Management*, Vol. 40 No. 11, pp. 842-864.
9. Ducret, R. (2014), "Parcel deliveries and urban logistics: changes and challenges in the courier express and parcel sector in Europe—the French case", *Research in Transportation Business and Management*, Vol. 11, pp. 15-22. Modelling e-commerce order arrival
10. Ekici, B.B. and Aksoy, U.T. (2011), "Prediction of building energy needs in early stage of design by using ANFIS", *Expert Systems with Applications*, Vol. 38 No. 5, pp. 5352-5358.
11. GëuNeri, A.F., Ertay, T. and YëuCel, A. (2011), "An approach based on ANFIS input selection and Modelling for supplier selection problem", *Expert Systems with Applications*, Vol. 38 No. 12, pp. 14907-14917.
12. Guresen, E., Kayakutlu, G. and Daim, T.U. (2011), "Using artificial neural network models in stock market index prediction", *Expert Systems with Applications*, Vol. 38 No. 8, pp. 10389-10397.
13. Hëubner, A., Kuhn, H. and Wollenburg, J. (2016), "Last mile fulfilment and distribution in omni-channel grocery retailing: a strategic planning framework", *International Journal of Retail and Distribution Management*, Vol. 44 No. 3, pp. 228-247.
14. Jang, J.S. (1993), "ANFIS: adaptive-network-based fuzzy inference system", *IEEE transactions on systems, man, and cybernetics*, Vol. 23 No. 3, pp. 665-685.
15. Jassbi, J., Seyedhosseini, S.M. and Pilevari, N. (2010), "An adaptive neuro fuzzy inference system for supply chain agility evaluation", *International Journal of Industrial Engineering and Production Research*, Vol. 20 No. 4, pp. 187-196.
16. Kar, S., Das, S. and Ghosh, P.K. (2014), "Applications of neuro fuzzy systems: a brief review and future outline", *Applied Soft Computing*, Vol. 15, pp. 243-259.
17. Karaboga, D. and Kaya, E. (2019), "Adaptive network based fuzzy inference system (ANFIS) training approaches: a comprehensive survey", *Artificial Intelligence Review*, Vol. 52 No. 4, pp. 2263-2293.
18. Klapp, M.A., Erera, A.L. and Toriello, A. (2018), "The dynamic dispatch waves problem for same-day delivery", *European Journal of Operational Research*, Vol. 271 No. 2, pp. 519-534.
19. Lang, G. and Bressolles, G. (2013), "Economic performance and customer expectation in e-Fulfilment systems: a multi-channel retailer perspective", *Supply Chain Forum: International Journal*, Vol. 14 No. 1, pp. 16-26, Taylor & Francis.
20. Leung, K.H., Choy, K.L., Siu, P.K., Ho, G.T., Lam, H.Y. and Lee, C.K. (2018), "A B2C e-commerce intelligent system for re-engineering the e-order fulfilment process", *Expert Systems with Applications*, Vol. 91, pp. 386-401.
21. Lim, S.F.W., Jin, X. and Srari, J.S. (2018), "Consumer-driven e-commerce: a literature review, design framework, and research agenda on last-mile logistics models", *International Journal of Physical Distribution and Logistics Management*, Vol. 48 No. 3, pp. 308-332.
22. Lu, W., McFarlane, D., Giannikas, V. and Zhang, Q. (2016), "An algorithm for dynamic order-picking in warehouse operations", *European Journal of Operational Research*, Vol. 248 No. 1, pp. 107-122.
23. Ma, Y., Wang, N., Che, A., Huang, Y. and Xu, J. (2013), "The bullwhip effect on product orders and inventory: a perspective of demand forecasting techniques", *International Journal of Production Research*, Vol. 51 No. 1, pp. 281-302.
24. MacCarthy, B.L., Zhang, L. and Muyltermans, L. (2019), "Best performance frontiers for buy-onlinepickup- in-store order fulfilment", *International Journal of Production Economics*, Vol. 211, pp. 251-264.
25. Mangiaracina, R., Marchet, G., Perotti, S. and Tumino, A. (2015), "A review of the environmental implications of B2C e-commerce: a logistics perspective", *International Journal of Physical Distribution and Logistics Management*, Vol. 45 No. 6, pp. 565-591.
26. Masmoudi, M., Benaissa, M. and Chabchoub, H. (2014), "Optimisation of e-commerce logistics distribution system: problem modelling and

- exact resolution", *International Journal of Business Performance and Supply Chain Modelling*, Vol. 6 Nos 3-4, pp. 358-375.
27. Nguyen, D.H., de Leeuw, S. and Dullaert, W.E. (2018), "Consumer behaviour and order fulfilment in online retailing: a systematic review", *International Journal of Management Reviews*, Vol. 20 No. 2, pp. 255-276.
28. Ramanathan, R. (2010), "The moderating roles of risk and efficiency on the relationship between logistics performance and customer loyalty in e-commerce", *Transportation Research Part E: Logistics and Transportation Review*, Vol. 46 No. 6, pp. 950-962.

AUTHOR PROFILE



Hussein Yehea Alsteif, Girne American University, Girne, North Cyprus, E-mail: hsteif@gmail.com

Skills and Interests

Degree	Subject	University	Year
Bachelors	Information Systems & Software Engineering	Damascus university	2004
Masters	Management Information Systems	The Arab academy for banking and financial science	2010
PhD	Management Information Systems	Girne American university	Under thesis monitoring



Prof. Dr. Murat Akkaya, Girne American University, Girne, North Cyprus.

Skills and Interests

Degree	Subject	University	Year
BSC	Computer Studying Information Tech	Eastern Mediterranean University	2002
MSC	Computer Engineering	Girne American University	2004
PHd	Management Information Systems	Girne American University	2013

