Introduction to the Special Issue on Neural Networks in Financial Engineering

THERE are several phases that an emerging field goes through before it reaches maturity, and computational finance is no exception. There is usually a trigger for the birth of the field. In our case, new techniques such as neural networks, significant progress in computing technology, and the need for results that rely on more realistic assumptions inspired new researchers to revisit the traditional problems of finance, problems that have often been tackled by introducing simplifying assumptions in the past. The result has been a wealth of new approaches to these time-honored problems, with significant improvements in many cases.

After the initial excitement settles in, the crest breaking question is whether this new community of researchers can produce sufficient results to sustain the field, and whether practitioners will find these results to be of quality, novelty, and relevance to make a real impact. Successful applications of computational methods to a wide variety of new and emerging problems such as complicated portfolio assignment and optimization problems, prediction, and pricing of complex derivatives rang the bell signifying the entry of computational finance as an established field. Further down the road, there comes a point at which the reality of the field is no longer in dispute, and one might even debate cosmetic issues such as what the "proper" name of the field should be, who belongs and who does not. We are now at this (enviable?) stage with computational finance.

Indeed, the current field may owe its origins to neural networks. In 1993, the first conference on "Neural Networks in the Capital Markets" (NNCM) was held in London, and was immensely popular. The interest from both academicians and practitioners made it clear that a new field was emerging. NNCM continued to alternate between England and the United States, and in 1997 changed its name to the International Conference on Computational Finance, reflecting the broadening of the technologies that started with neural networks. The IEEE made a major contribution to the field with the CIFEr conference (Computational Intelligence in Financial Engineering) that started in New York in 1995. This conference attracted Wall Street to the research results of the academic community, and helped consolidate the new field. Recently also, new journals and other conferences started coming to the scene. A number of major universities now have a computational finance program as part of their curriculum.

With this special issue, we hope to get the word out to a larger audience that computational finance is happening, and that its core science and technology are accessible to engineers and computer scientists, especially those whose interests coincide

with these TRANSACTIONS. We have made a conscious effort to sustain the quality, variety, and accessibility of the papers that appear in this issue. We believe that this collection of papers on a wide spectrum of topics comes with a high signal-to-noise ratio for a reader that wants to learn about the state of the art in this field. Some of the topics covered are theory, derivatives pricing techniques, agent-based market simulation and analysis, time series prediction, interest rates and model calibration, machine learning techniques, volatility prediction, portfolio analysis, stock selection, and credit assessment and bankruptcy prediction.

In many financial applications of neural networks a large number of candidate inputs is typically available. One is faced with the problem of choosing a set of inputs that are effective for the particular application and for the particular neural-network model employed. In the first paper, White and Racine propose new statistical tests that would determine the relevance of an input or a group of inputs.

The way neural networks are employed for time series applications is via a nonlinear autoregressive model with exogenous inputs. The paper by Chen *et al.* considers such a model using three classes of neural networks. They prove tight root mean square (rms) bounds on the convergence rates of the neural-network estimators. They also apply the models to the problem of forecasting inflation.

The concept of risk-neutrality is a fundamental concept in the theory of finance, forming the basis for many simulation-based derivatives pricing techniques. The paper by Magdon-Ismail provides a derivation of this concept from first principles, and gives examples of how risk-neutrality can be employed for the pricing of different financial instruments.

For American style derivatives with high-dimensional underlying state space, even computational pricing approaches become very difficult to implement because of the computational complexity. Tsitsiklis and Van Roy propose a computational method based on dynamic programming. Specifically they use the concept of value iteration (also called Q-learning). They prove convergence of their method.

While analytical solutions exist for the problem of pricing of European style options and American call options, the analytical pricing of the American put option has defied the research community for almost three decades. Keber and Schuster consider such a problem, and they apply the genetic programming approach to the pricing of the American put vega (sensitivity to volatility). By using a search space as the "set of all analytical expressions" they arrive at an accurate approximation.

The concept of risk neutral density allows pricing of derivatives by taking expectations. Schittenkopf and Dorffner consider

such a problem, and present a new approach for risk neutral density extraction from options prices. Specifically they use an extension of the concept of mixture density networks. The model is applied to a large data base for FTSE 100 intraday options data.

Neural networks have been shown to successfully learn to estimate a pricing formula for options, with good out of sample performance. Gencay and Qi consider the European call option, and compare between several regularization techniques, such as cross validation, Bayesian regularization, early stopping, bagging, as well as the standard Black Scholes benchmark. They test these methods on S&P 500 call options.

A novel aproach to studying the dynamics of financial markets is to model the buyers and sellers as interacting agents. Out of this microlevel interaction emerges the price dynamics. Zimmermann et al. use this approach, and propose a perceptron model for each micro-agent. The agregate market then acts like a neural network which can be trained on prior data to extract the dynamics. They use their multiagent representation to model and hence forecast multiple foreign exchange markets.

There has been a long debate in the literature whether certain financial markets at certain time horizons are predictable, but most studies tested linear prediction models or standard technical indicators. Dempster et al. consider the intraday foreign exchange market, specifically the US/GBP market. They develop, test, and compare between a number of AI-based trading techniques: genetic algorithm, reinforcement learning, and Markov chain-based method.

Along the lines of testing the predictability and the linearity/nonlinearity of particular markets, Madeiros et al. consider 14 monthly foreign exchange time series. By testing a multilayer network and a novel model against benchmarks they reject the null hypothesis of linearity in ten out of the 14 series.

Predicting financial time series over different time scales can give different forecasts. Zhang et al. propose a model where they combine forecasts of different time scales to obtain a superior forecast. They use the wavelet transform, as well as a neural network predictor for such purpose. They apply their method to the ten-year treasury bond futures and the Australian dollar futures markets.

The paper by Ingber and Mondescu presents a physics-related derivation of financial market indicators based on a path integral approach to the conditional probability density transition. They combine these indicators with a trading system to provide a full end-to-end trading system which they apply to the S&P 500 futures market.

The paper by Abu-Mostafa deals with the calibration of the Vasicek interest-rate model using consistency hints. It shows how the learning from hints technique, which has been traditionally applied to neural-network models, can be successfully applied to financial models.

The paper by Van Gestel et al. provides a combination of the least square support vector machine and the evidence framework for time series and volatility modeling via three levels of inference. The authors present an iterative implementation of the theory and apply the method to the prediction of the weekly T-bill rate and the daily DAX30.

Xu presents a unification of various paradigms, such as PCA, ICA, linear regression, HMM, and factor analysis. These techniques are of much use in the machine learning community, and they can also be quite useful in finance as well. The author shows how the well-known arbitrage pricing theory can be posed in the proposed framework.

For time series prediction models there are several statistical tests for checking model adequacy. Among them is the autocorrelation test for the residuals. Refenes and Holt generalize the standard Durbin-Watson autocorrelation test to the neural regression problem. They develop a neural-network-based method for forecasting the implied volatility for a high-frequency time series for the Spanish stock index, and apply their model adequacy method.

Tino et al. consider the problem of trading the volatility time series. They consider a time series of straddles on the financial indexes DAX and FTSE 100. They use recurrent neural networks, and compare it with a fixed-order Markov model and GARCH. Then, they propose a combination of recurrent networks and Markov model that leads to improvement in prediction accuracy.

Developing trading systems and optimizing portfolios is a complex task that can be posed as a stochastic control problem. By taking this view, Moody and Saffell develop a direct reinforcement method called recurrent reinforcement learning. Their algorithm maximizes the risk-adjusted return in a dynamic fashion, while taking into account transaction costs. They apply their method to high-frequency foreign exchange trading and stock index/T-bill portfolio system.

Chapados and Bengio consider also the portfolio optimization problem. They devise a method based on active control of the value-at-risk of the portfolio. They develop a backpropagation-type algorithm that trains a network to directly make the portfolio allocation decisions. They apply their method to a sector trading system for Canadian stocks.

Stock analysis has been one of the important applications of neural networks. It has been well-documented in the finance literature that there is predictability in earnings surprises. However, these studies used linear models. Chou and Dhar consider this problem, and develop nonlinear forecasting models for earnings surprises and market returns. They compare between tree induction algorithms, neural networks, genetic algorithms, naive Bayes method, and linear regression.

Chan and Genovese consider the stock selection problem. Using financial statement and other indicators, they apply and compare between refined linear model, additive nonlinear model, and neural networks. They apply the methods on the Russell 1000 stocks.

Predicting bankruptcies for corporations is one of the earliest financial applications of neural networks. Atiya presents a survey of such application. In addition, he develops a default prediction model and proposes new indicators, that lead to enhancing the classification performance.

In high-dimensional problems, visualization of the problem can shed much light on the qualitative relationships between the variables. Kaski *et al.* consider the self-organizing map approach, and propose a new metric based on the Fisher information matrix. They apply their method to the bankruptcy prediction problem.

YASER S. ABU-MOSTAFA, *Guest Editor*
California Institute of Technology
Departments of Electrical Engineeering, Computer Science, and
Computation and Neural Systems
Pasadena, CA 91125 USA

AMIR F. ATIYA, *Guest Editor*
California Institute of Technology
Department of Electrical Engineering
Pasadena, CA 91125 USA

MALIK MAGDON-ISMAIL, *Guest Editor*
Rensselaer Polytechnic Institute
Department of Computer Science
Troy, NY 12180 USA

HALBERT WHITE, *Guest Editor*
University of California
Department of Economics
San Diego, CA 92093 USA

Yaser S. Abu-Mostafa received the B.Sc. degree from Cairo University, Egypt, in 1979, the M.S.E.E. degree from Georgia Tech, Atlanta, in 1981, and the Ph.D. degree from the California Institute of Technology (Caltech), Pasadena, in 1983.

He is Professor of Electrical Engineering and Computer Science, and Head of the Learning Systems Group, at the California Institute of Technology. His research interests are in the areas of machine learning, computational finance, and neural networks, and he has more than 80 technical publications including two articles in *Scientific American.*

Dr. Abu-Mostafa was a founding member of the IEEE Neural Networks Council, and the founding program chairman of the Neural Information Processing Systems (NIPS) conference. He chaired a number of national and international conferences, most recently the international conference on *Computational Finance* (CF'99), and has served on the boards of ten journals. Since 1988, he has been a Technical Consultant for a number of financial firms, including Citibank for nine years. Among his awards are the 1996 Richard P. Feynman Prize for excellence in teaching, and the 1998 Kuwait State Award in Applied Science. He was awarded the Clauser Prize from Caltech for the most original doctoral thesis.

Amir F. Atiya (S'86–M'90–SM'97) received the B.S. degree in 1982 from Cairo University, Cairo, Egypt, and the M.S. and Ph.D. degrees in 1986 and 1991 from California Intsitute of Technology (Caltech), Pasadena, all in electrical engineering

He held a Research Associate position at Texas A&M University, and from 1991 to 1993, he was a Senior Research Scientist at QANTXX, in Houston, TX, a financial modeling firm. In 1993, he joined the Computer Engineering Department at Cairo University as an Assistant Professor. Since March 1997, he has been a Visiting Associate in the Department of Electrical Engineering at Caltech. In addition, from January 1998 to January 2000, he was a Research Scientist at Simplex Risk Management Inc., Hong Kong. He has been an Associate Editor for IEEE TRANSACTIONS ON NEURAL NETWORKS, since 1998. His research interests are in the areas of neural networks, learning theory, computational finance, financial forecasting, pattern recognition, Monte Carlo methods, and optimization theory. He has also been active in developing quantitative and AI-based financial market trading systems.

Dr. Atiya received the highly regarded Egyptian State Prize for Best Research in Science and Engineering, in 1994. He also received the Young Investigator Award from the International Neural Network Society in 1996. He served on the organizing and program committees of several conferences, the most recent of which was Computational Finance CF2000, London.

Malik Magdon-Ismail (S'98–M'99) received the B.S. degree in physics from Yale University, New Haven, CT, in 1993 and the Masters degree in physics and the Ph.D. degree in electrical engineering with a minor in physics from the California Institute of Technology, Pasadena, in 1995 and 1998, respectively.

He has been a Research Fellow in the Learning Systems Group at Caltech from 1998 to 2000, and is currently an Assistant Professor of Computer Science at Rensselaer Polytechnic Institute (RPI), Troy, NY. He is a founding faculty member of the Machine and Computational Learning Group (MaCoL) at RPI. His research interests include the theory and applications of machine and computational learning (supervised, reinforcement, and unsupervised), computational finance and bioinformatics, and, he has numerous publications in refereed journals and conferences, has been a financial consultant, and has collaborated with a number of companies.
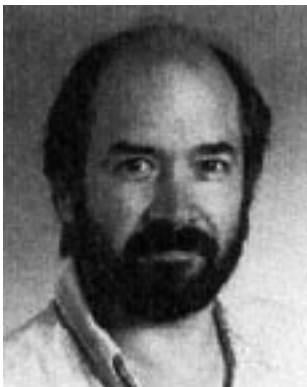
Halbert White (A'93–SM'99) received the Ph.D. degree from the Massachusetts Institute of Technology, Cambridge, in 1976.

He began his career as Assistant Professor of Economics at the University of Rochester, after which he assumed his current position (now Professor Above Scale) at the University of California, San Diego. His interests include econometrics, statistics, artificial neural networks, forecasting, and financial markets.

Dr. White has been a Fellow of the Econometric Society since 1983, received a Guggenheim Fellowship for 1988 to 1989 for the study of artificial neural networks, and was elected a Fellow of the American Academy of Arts and Sciences in 1999.