Preface to the First Edition

This book evolved from the lecture notes of a first-year graduate course entitled "Neural Networks" which I taught at the Department of Electrical Engineering and Computer Science of the University of Illinois at Chicago over the years 1990–1996. Whereas that course was a first-year graduate course, several Senior-Year undergraduate students from different engineering departments, attended it with little difficulty. It was mainly for historical and scheduling reasons that the course was a graduate course, since no such course existed in our program of studies and in the curricula of most U.S. universities in the Senior Year Undergraduate program. I therefore consider this book, which closely follows these lecture notes, to be suitable for such undergraduate students. Furthermore, it should be applicable to students at that level from essentially every science and engineering University department. Its prerequisites are the mathematical fundamentals in terms of some linear algebra and calculus, and computational programming skills (not limited to a particular programming language) that all such students possess.

Indeed, I strongly believe that Neural Networks are a field of both intellectual interest and practical value to all such students and young professionals. Artificial neural networks not only provide an understanding into an important computational architecture and methodology, but they also provide an understanding (very simplified, of course) of the mechanism of the biological neural network.

Neural networks were until recently considered as a "toy" by many computer engineers and business executives. This was probably somewhat justified in the past, since neural nets could at best apply to small memories that were analyzable just as successfully by other computational tools. I believe (and I tried in the later chapters below to give some demonstration to support this belief) that neural networks are indeed a valid, and presently, the only efficient tool, to deal with very large memories.

The beauty of such nets is that they can allow and will in the near-future allow, for instance, a computer user to overcome slight errors in representation, in programming (missing a trivial but essential command such as a period or any other symbol or character) and yet have the computer execute the command. This will obviously require a neural network buffer between the keyboard and the main pro-

grams. It should allow browsing through the Internet with both fun and efficiency. Advances in VLSI realizations of neural networks should allow in the coming years many concrete applications in control, communications and medical devices, including in artificial limbs and organs and in neural prostheses, such as neuromuscular stimulation aids in certain paralysis situations.

For me as a teacher, it was remarkable to see how students with no background in signal processing or pattern recognition could easily, a few weeks (10–15 hours) into the course, solve speech recognition, character identification and parameter estimation problems as in the case studies included in the text. Such computational capabilities make it clear to me that the merit in the neural network tool is huge. In any other class, students might need to spend many more hours in performing such tasks and will spend so much more computing time. Note that my students used only PCs for these tasks (for simulating all the networks concerned). Since the building blocks of neural nets are so simple, this becomes possible. And this simplicity is the main feature of neural networks: A house fly does not, to the best of my knowledge, use advanced calculus to recognize a pattern (food, danger), nor does its CNS computer work in picosecond-cycle times. Researches into neural networks try, therefore, to find out why this is so. This leads and led to neural network theory and development, and is the guiding light to be followed in this exciting field.

Daniel Graupe
Chicago, IL
January 1997

Preface to the Second Edition

The Second Edition contains certain changes and additions to the First Edition. Apart from corrections of typos and insertion of minor additional details that I considered to be helpful to the reader, I decided to interchange the order of Chapters 4 and 5 and to rewrite Chapter 13 so as to make it easier to apply the LAMSTAR neural network to practical applications. I also moved the Case Study 6.D to become Case Study 4.A, since it is essentially a Perceptron solution.

I consider the Case Studies important to a reader who wishes to see a concrete application of the neural networks considered in the text, including a complete source code for that particular application with explanations on organizing that application. Therefore, I replaced some of the older Case Studies with new ones with more detail and using most current coding languages (MATLAB, Java, C++). To allow better comparison between the various neural network architectures regarding performance, robustness and programming effort, all Chapters dealing with major networks have a Case Study to solve the same problem, namely, character recognition. Consequently, the Case studies 5.A (previously, 4.A, since the order of these chapters is interchanged), 6.A (previously, 6.C), 7.A, 8.A, have all been replaced with new and more detailed Case Studies, all on character recognition in a 6×6 grid. Case Studies on the same problem have been added to Chapter 9, 12 and 13 as Case Studies 9.A, 12.A and 13.A (the old Case Studies 9.A and 13.A now became 9.B and 13.B). Also, a Case Study 7.B on applying the Hopfield Network to the well known Traveling Salesman Problem (TSP) was added to Chapter 7. Other Case Studies remained as in the First Edition.

I hope that these updates will add to the readers' ability to better understand what Neural Networks can do, how they are applied and what the differences are between the different major architectures. I feel that this and the case studies with their source codes and the respective code-design details will help to fill a gap in the literature available to a graduate student or to an advanced undergraduate Senior who is interested to study artificial neural networks or to apply them.

Above all, the text should enable the reader to grasp the very broad range of problems to which neural networks are applicable, especially those that defy analysis and/or are very complex, such as in medicine or finance. It (and its Case Studies)

should also help the reader to understand that this is both doable and rather easily programmable and executable.

Daniel Graupe
Chicago, IL
September 2006