Open access • Proceedings Article • DOI:10.1109/IJCNN.2003.1223937

Training support vector machines with particle swarms — Source link

U. Paquet, Andries P. Engelbrecht

Institutions: University of Pretoria

Published on: 20 Jul 2003 - International Joint Conference on Neural Network

Topics: Sequential minimal optimization, MNIST database, Support vector machine, Quadratic programming and Particle swarm optimization

Training Support Vector Machines with Particle Swarms

deur

Ulrich Paquet

Voorgelê as 'n deel van die vereistes vir die graad

Magister Scientiae

in die Fakulteit Ingenieurswese, Bouomgewing, en Inligtingstegnologie

Universiteit van Pretoria

November 2003

Die finansiële ondersteuning van die *National Research Foundation* (NRF) tot hierdie navorsing word hierdeur erken. Opinies en gevolgtrekkings in hierdie verhandeling is dié van die outeur, en kan nie noodwendig aan die NRF toegeskryf word nie.

Training Support Vector Machines with Particle Swarms

by

Ulrich Paquet

Submitted in partial fulfilment of the requirements for the degree

Magister Scientiae

in the Faculty of Engineering, Built Environment and Information Technology

University of Pretoria

November 2003

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed in this thesis and conclusions arrived at, are those of the author and not necessarily to be attributed to the National Research Foundation.

Training Support Vector Machines with Particle Swarms

deur

Ulrich Paquet

Opsomming

Deelswerms kan met gemak gebruik word om 'n funksie, wat beperk word deur 'n stel lineêre beperkings, te optimeer. 'n "Lineêre Deelswermoptimeerder" en 'n "Konvergente Lineêre Deelswermoptimeerder" word ontwikkel om sulke beperkte funksies te optimeer. As die hele swerm aanvanklik slegs uit geldige oplossings bestaan, dan kan die swerm die beperkte funksie optimeer sonder om ooit weer die stel beperkings te oorweeg. Die Konvergente Lineêre Deelswermoptimeerder" oorkom die waarskynlikheid van vroegtydige konvergensie, wat deur die Lineêre Deelswermoptimeerder" getoon word. Om 'n Steunvektormasjien te leer moet 'n beperkte kwadratiese programmeringsprobleem opgelos word, en die Konvergente Lineêre Deelswermoptimeerder voldoen aan die behoeftes van 'n optimeringsmetode vir Steunvektormasjiene. Deelswerms is intuïtief en maklik om te implementeer, en word aangebied as 'n alternatief tot huidige metodes om Steunvektormasjiene te leer.

Studieleier: Prof. A.P. Engelbrecht
Departement Rekenaarwetenskap
Degree: Magister Scientiae

Training Support Vector Machines with Particle Swarms

by

Ulrich Paquet

Abstract

Particle swarms can easily be used to optimise a function with a set of linear equality constraints, by restricting the swarm's movement to the constrained search space. A "Linear Particle Swarm Optimiser" and "Converging Linear Particle Swarm Optimiser" is developed to optimise linear equality-constrained functions. It is shown that if the entire swarm of particles is initialised to consist of only feasible solutions, then the swarm can optimise the constrained objective function without ever again considering the set of constraints. The Converging Linear Particle Swarm Optimiser overcomes the Linear Particle Swarm Optimiser's possibility of premature convergence. Training a Support Vector Machine requires solving a constrained quadratic programming problem, and the Converging Linear Particle Swarm Optimiser ideally fits the needs of an optimisation method for Support Vector Machine training. Particle swarms are intuitive and easy to implement, and is presented as an alternative to current numeric Support Vector Machine training methods.

Supervisor: Prof. A.P. Engelbrecht
Department of Computer Science
Degree: Magister Scientiae

Preface

The question that originally spurred the research in this thesis was - "can a Particle Swarm Optimiser be used to train a Support Vector Machine, and to what extent will it be successful?"

Training a Support Vector Machine (SVM) involves solving a quadratic programming problem, with a single linear constraint, and a set of non-negativity constraints. At first this problem seemed trivial - the objective function that needs to be minimised is convex, and the Particle Swarm Optimiser (PSO) will not be trapped in any local minima.

The difficulty in the problem arose with developing a method to handle the linear constraint. This has led to the development of the Linear (and Converging Linear) PSO algorithms (LPSO and CLPSO), which both have unique properties needed not only for handling the single linear constraint, but any set of (feasible) linear constraints. The non-negativity constraints have led to the extension of both new Particle Swarm algorithms to include cases when constraints appear as boxed constraints. With the addition of slack variables to an optimisation problem with linear constraints, it becomes possible to solve any of these problems.

The main contributions made by this thesis are therefore:

1. An algorithm for SVM training, which is based on analysis of a method for decomposing the SVM quadratic programming problem.
2. The development of LPSO for general optimisation problems, and a proof of a condition on the initial swarm guaranteeing that any point in the search space can be reached.
3. A proof that LPSO is ideally suited for linearly constrained optimisation, with a precondition needed for LPSO to always satisfy linear equality constraints.
4. The extension of LPSO to CLPSO to preclude premature convergence, and a proof that CLPSO will at least converge to a local minimum.
5. The addition of a method to LPSO and CLPSO needed for inequality constraint handling, and the implementation of CLPSO with inequality constraints as an optimiser in SVM training.

In a sense this thesis has delivered more than its original aim. The new PSO algorithms will probably be of greater importance to further milestones in the Swarm Intelligence community, than its application in SVM training.

Chapter 1 puts SVMs under the magnifying glass, and sets the main optimisation problem (a quadratic programming problem) that forms the backbone of this thesis. SVM training has unique problems of its own, primarily because the training problem shows quadratic growth as the training set size increases. Methods for SVM training are discussed in *Chapter 2*, and a training algorithm, based on standard methods of decomposing the main optimisation problem into subproblems, are used to construct a correct training algorithm. *Chapter 3* introduces PSO as an optimisation algorithm, and discusses some of the recent advancements to the PSO method itself. The PSO is extended to handle constrained problems in *Chapter 4*, and LPSO and CLPSO are developed. This extension includes a rigorous analysis of the newly developed algorithms. The successes and failures of LPSO and CLPSO are empirically shown in *Chapter 5*. It is also shown how the CLPSO can be used to train SVMs from a very large character recognition dataset. Finally, *Chapter 6* provides an overview, and gives some thoughts for further research.

Many people have contributed to the successful completion of this thesis. Foremost, I am greatly indebted to professor Andries Engelbrecht for introducing me to the world of artificial intelligence, and for his patient guidance throughout my research.

Ulrich Paquet
Pretoria, South Africa
June 2003

Commit thy works unto the LORD, and thy thoughts shall be established. Proverbs 16:3

Contents

List of Figures

List of Tables