



Novel evolutionary-based methods for the robust training of SVR and HMDH regressors [

Gascón Moreno, Jorge

2016

Tesis y escritos académicos

Monografía

En los últimos años se han consolidado una serie de diferentes métodos y algoritmos para problemas de aprendizaje máquina y optimización de sistemas, que han dado lugar a toda una corriente de investigación conocida como Soft-Computing. El término de Soft-Computing hace referencia a una colección de técnicas computacionales que intenta estudiar, modelar y analizar fenómenos muy complejos, para los que los métodos convencionales no proporcionan soluciones completas, o no las proporcionan en un tiempo razonable. Dentro de lo que se considera como Soft-Computing existen una gran cantidad de técnicas tales como Redes Neuronales, Máquinas de Vectores Soporte (SVM), Redes Bayesianas, Computación Evolutiva (Algoritmos Genéticos, Algoritmos Evolutivos etc), etc. La investigación de la Tesis está enfocada en dos de estas técnicas, en primer lugar las máquinas de vectores soporte de regresión (SVR) y en segundo lugar a las GMDH (Group Method of Data Handling). Las SVM son una técnica ideada por Vapnik, basada en el principio de minimización del riesgo estructural y la teoría de los métodos kernel, que a partir de un conjunto de datos construye una regla de decisión con la cual intentar predecir nuevos valores para dicho proceso a partir de nuevas entradas. La eficiencia de los sistemas SVM ha hecho que tengan un desarrollo muy significativo en los últimos años y se hayan utilizado en una gran cantidad de aplicaciones tanto para clasificación como para problemas de regresión (SVR). Uno de los principales problemas es la búsqueda de los que se conoce como hiper-parámetros. Estos parámetros no pueden ser calculados de forma exacta, por lo que se hace necesario testear un gran número de combinaciones, para obtener unos parámetros que generen una buena función de estimación. Debido a esto el tiempo de entrenamiento suele ser elevado y no siempre los parámetros encontrados generan una buena solución: ya sea porque el algoritmo de búsqueda tenga un pobre rendimiento o porque el modelo generado está sobre-entrenado. En esta Tesis se ha desarrollado un nuevo algoritmo de tipo evolutivo para el entrenamiento con kernel multi-paramétrico. Este nuevo algoritmo tiene en cuenta un parámetro - distinto, para cada una de las dimensiones del espacio de entradas. En este caso, debido al incremento del número de parámetros no puede utilizarse una búsqueda en grid clásica, debido al coste computacional que conllevaría. Por ello, en esta Tesis se propone la utilización de un algoritmo evolutivo para la obtención de los valores óptimos de los parámetros de la SVR y la aplicación de nuevas cotas para los parámetros de este kernel multi-paramétrico. Junto con esto, se han desarrollado nuevos métodos de validación que mejoren el rendimiento de las técnicas de regresión en problemas data-driven. La idea es obtener mejores modelos en la fase de entrenamiento del algoritmo, de tal forma que el desempeño con el conjunto de test

mejore, principalmente en lo que a tiempo de entrenamiento se refiere y en el rendimiento general del sistema, con respecto a otros métodos de validación clásicos como son K-Fold cross-validation, etc. El otro foco de investigación de esta Tesis se encuentra en la técnica GMDH, ideada en los años 70 por Ivakhnenko. Es un método particularmente útil para problemas que requieran bajos tiempos de entrenamiento. Es un algoritmo auto-organizado, donde el modelo se genera de forma adaptativa a partir de los datos, creciendo con el tiempo en complejidad y ajustándose al problema en cuestión, hasta que el modelo alcanza un grado de complejidad óptima, es decir, no es demasiado simple ni demasiado complejo. De esta forma el algoritmo construye el modelo en base a los datos de los que dispone y no a una idea preconcebida del investigador, como ocurre en la mayoría de las técnicas de Soft-Computing. Las GMDH también tienen algunos inconvenientes como son los errores debido al sobre-entrenamiento y la multicolinealidad, esto hace que en algunas ocasiones el error sea elevado si lo comparamos con otras técnicas. Esta Tesis propone un nuevo algoritmo de construcción de estas redes basado en un algoritmo de tipo hiper-heurístico. Esta aproximación es un concepto nuevo relacionado con la computación evolutiva, que codifica varios heurísticos que pueden ser utilizados de forma secuencial para resolver un problema de optimización. En nuestro caso particular, varios heurísticos básicos se codifican en un algoritmo evolutivo, para crear una solución hiper-heurística que permita construir redes GMDH robustas en problemas de regresión. Todas las propuestas y métodos desarrollados en esta Tesis han sido evaluados experimentalmente en problemas benchmark, así como en aplicaciones de regresión reales

<https://rebiunoda.pro.baratznet.cloud:38443/OpacDiscovery/public/catalog/detail/b2FpOmNlbGVicmF0aW9uOmVzLmJhcmF0ei5yZW4vMjc3OTMyNDk>

Título: Novel evolutionary-based methods for the robust training of SVR and HMDH regressors [Recurso electrónico] Jorge Gascón Moreno ; dirs., Sancho Salcedo Sanz, José Antonio Portilla Figueras ; Universidad de Alcalá, Escuela Politécnica Superior, Departamento de Teoría de la Señal y Comunicaciones

Editorial: 2016

Descripción física: 1 disco compacto (CD-ROM)

Tesis: Tesis Univ. Alcalá, 2016

Bibliografía: Bibliografía: p. [131]-142

Tipo recurso electrónico: Texto

Formato físico adicional: También disponible en TESEO

Lengua: Texto en inglés y resumen en inglés y español

Materia: Proceso de vectores (Informática) Algoritmos de tratamiento agrupado de datos. Soft Computing Computación evolutiva Algoritmos heurísticos. Inteligencia artificial Señales, Teoría de (Telecomunicación)

Autores: Salcedo Sanz, Sancho Portilla Figueras, José Antonio

Entidades: Universidad de Alcalá de Henares. Departamento de Teoría de la Señal y Comunicaciones

Baratz Innovación Documental

- Gran Vía, 59 28013 Madrid
- (+34) 91 456 03 60
- informa@baratz.es