

Modelos de Aprendizaje Supervisados: aplicaciones para la predicción de incendios forestales en la provincia de Córdoba

Marina Cardenas¹, Ricardo Medel¹, Julio Castillo¹, Juan C. Vázquez¹, Osvaldo Casco¹

¹ Laboratorio de Investigación de Software/Dpto. Ingeniería en Sistemas de Información/ Facultad Regional Córdoba/ Universidad Tecnológica Nacional
{ ing.marinacardenas, ricardo.h.medel, jotacastillo, jcjvazquez, casqui.159}@gmail.com

Resumen

Mediante el empleo de técnicas de aprendizaje automático, tales como las redes neuronales o SVM (Máquinas de Vectores de Soporte), queremos capturar la presencia de patrones de comportamiento en los factores que influyen en la ocurrencia de incendios forestales, ya sean humanos o de índole climática.

El objetivo de este proyecto es elaborar un modelo que permita pronosticar la ocurrencia de incendios forestales en la Provincia de Córdoba, especialmente en las sierras de Córdoba y la región del Parque Chaqueño de la provincia.

Cabe destacar que el modelo construido servirá no solo para pronosticar la ocurrencia de un incendio en una determinada zona, sino también brindará información sobre la cantidad de hectáreas afectadas por el mismo.

Actualmente estamos realizando pruebas con datos recopilados en artículos periodísticos, disponibles en Internet, y bases de datos de estaciones meteorológicas situadas en la provincia.

Palabras clave: Pronósticos, modelo, incendios, forestales, meteorología, predicción.

Contexto

El presente proyecto se encuentra consolidado dentro de la línea de investigación que se dedica a la resolución de problemas físicos, reales y sociales, a través del uso de herramientas computacionales basadas en aprendizaje automático.

Este proyecto se lleva a cabo en el Laboratorio de Investigación de Software del Dpto. de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional, Facultad Regional Córdoba (UTN-FRC), siendo acreditado por la Secretaría de Ciencia, Tecnología y Posgrado de la UTN.

Nuestro trabajo se enmarca dentro del Grupo GIA (Grupo de Inteligencia Artificial) de la UTN-FRC, el cual tiene como objetivo general el investigar técnicas, algoritmos de inteligencia artificial, entre los que se destacan el estudio de las redes neuronales, autómatas celulares, análisis y procesamiento de imágenes, minería de datos, y su aplicabilidad y resolución de problemas de las ciencias naturales y de las ciencias sociales. El grupo está integrado por doctores, ingenieros, licenciados, becarios y pasantes.

De esta manera, se puede observar que se investigan técnicas de IA (Inteligencia

Artificial) tanto desde el punto de vista teórico, como desde el punto de vista práctico.

Introducción

Los incendios forestales son uno de los mayores problemas ambientales y producen un daño ecológico, económico y humano irreparable. Es por ello que una detección prematura de los mismos es una herramienta clave que puede permitir una lucha más eficaz contra este flagelo.

El objetivo de este trabajo es ofrecer una herramienta alternativa para apoyar las políticas de prevención de incendios forestales reduciendo la incertidumbre inherente a la naturaleza del problema planteado.

En la actualidad existen distintas aproximaciones que intentan dar soluciones empleando herramientas automáticas basadas en sensores locales (meteorológicos) [Chandler et. al., 1983], otras se basan en satélites [Rauste, 1996] y un tercer grupo de técnicas se basa en scanners de humo e infrarrojos [Ollero et. al., 2001]. En este proyecto pretendemos construir un sistema empleando Redes Neuronales Artificiales (RNA) [Cortez y Morais, 2007] y Máquinas de Vectores de Soporte (SVM- Support Vector Machines) [Hsu et. al., 2010] comunmente utilizadas para minería de datos [Cortez, 2010], y alimentarlas con datos meteorológicos no costosos, como los sugeridos por índices internacionales, con el objetivo de poder predecir la ocurrencia de un incendio forestal y la extensión del mismo. La información de salida que este sistema proveerá es de vital importancia para el planeamiento estratégico de los recursos destinados a combatir los incendios forestales.

Las técnicas propuestas son evaluadas respecto de su utilización con diversos tipos de datos de entrada, tales como los

componentes del índice FWI, datos espaciales, datos temporales, datos meteorológicos y datos satelitales provenientes de los sensores MODIS [Justicea y et. al., 2002] de los satélites Terra y Aqua.

Si bien el pronóstico de incendios es una tarea difícil, creemos que un enfoque basado en aprendizaje automático podrá inferir un modelo matemático útil para la planeación estratégica de recursos destinados a combatir los incendios forestales. Como resultado esperamos obtener un sistema que indique las áreas en las que ocurrirá un incendio y la cantidad de hectáreas que se verán afectadas.

Para la aplicación de técnicas de aprendizaje supervisado partimos de un conjunto de datos de iniciales (entradas) para las que se observa una serie de resultados (salidas) para una serie de eventos determinados. A partir de estos datos, es posible construir un modelo de predicción, que permitirá predecir los resultados que se obtendrán para un nuevo evento.

Este proceso de aprendizaje se realiza mediante un entrenamiento controlado por un agente externo (*supervisor*) que determina la respuesta que debería generar la red a partir de una entrada determinada. El supervisor comprueba la salida generada por el sistema y en el caso en que no coincida con la esperada procederá a modificar los pesos de las conexiones y/o su arquitectura con el objetivo de refinar y mejorar las predicciones del modelo. Este proceso se repite hasta alcanzar la configuración óptima.

A partir del análisis de las alternativas propuestas para la predicción asociada al dominio de problema planteado, distintos enfoques orientados a la predicción de incendios forestales han aplicado diversas técnicas de aprendizaje supervisado tales

como Redes Neuronales Artificiales (RNA) [Safi y Bouroumi, 2013] y [Vega Garcia et. al., 1999] y Máquinas de Vectores de Soporte (SVM) [Cortez y Morais, 2007] y [Kim, 2009]. Ambas aproximaciones utilizan principalmente datos meteorológicos y datos vinculados al índice FWI como entrada al modelo y obtienen resultados satisfactorios. Sin embargo, algunos autores sostienen que la mejor configuración del modelo utiliza una SVM con cuatro entradas, que son simplemente los datos meteorológicos de temperatura, humedad relativa, lluvia y viento [Cortez y Morais, 2007].

En el presente trabajo pretendemos implementar las técnicas de RNA y SVM con datos vinculados al FWI en forma conjunta con datos espaciales, temporales y satelitales provenientes de los sensores MODIS sobre alertas de focos de incendios de los satélites Terra y Aqua.

Líneas de Investigación, Desarrollo e Innovación

Como se mencionó inicialmente, el presente trabajo forma parte de la línea de investigación de Aprendizaje Automático, esto es, en la construcción de algoritmos que intentan inferir un modelo computacional con el fin de resolver un problema concreto, en este caso la predicción de incendios, problema que se caracteriza como complejo y altamente no lineal.

El presente proyecto se denomina *“Construcción de un modelo de pronósticos para predicción de incendios forestales en la provincia de Córdoba”* y fue aprobado por la Secretaria de Ciencia, Técnica y Posgrado de la UTN, comenzando formalmente en enero de 2013, como continuación del proyecto *“Modelado para la predicción de incendios forestales en la provincia de*

Córdoba” que finalizó en diciembre del 2012.

Resultados y Objetivos

Para poder realizar la etapa de entrenamiento requerida para las técnicas de aprendizaje supervisado es necesario contar con una serie de ejemplos de entradas, para lo cual se requiere contar con un registro de datos por cada una de las ellas.

Partiendo de los trabajos realizados por [Cortez y Morais, 2007] y [Safi y Bouroumi, 2013] se construyó el modelo a partir de los tipos de entradas utilizadas por ellos para probar el funcionamiento de este tipo de modelos con datos de la provincia de Córdoba, para luego proceder a la incorporación de nuevos tipos de datos basados en Índice Meteorológico de Riesgo de Incendios de Canadá (FWI) [Van Wagner y Pickett, 1985], que creemos ayudarán al mejoramiento de las predicciones.

Una vez definidos los datos que servirán como entradas y salida del modelo, se procedió a recopilar los datos vinculados a la Provincia de Córdoba para poder replicar el experimento que otros autores aplicaron al Parque Natural de Montesino, en Portugal [Safi y Bouroumi, 2013]. Sin embargo, se descubrió que tanto las estaciones de bomberos como las estaciones meteorológicas no llevaban a cabo ningún tipo de registro de incendios automatizado y que era necesario un sistema que permitiera capturar al menos estos datos. Para brindar una herramienta eficaz que nos permita realizar esta actividad, hemos desarrollado un sistema de soporte (SS) que permite la carga de la información relacionada a las mediciones diarias y a los siniestros forestales ocurridos en determinadas zonas de influencia de los cuarteles.

Además de los datos recopilados a través del sistema, nos encontramos en la fase de recopilación de datos satelitales provistos por el Sistema de Información de Incendios para la Gestión de Recursos (FIRMS, por sus siglas en inglés), que integra tecnologías de percepción remota y de Sistemas de Información Geográfica (SIG), para el envío de ubicaciones de incendios activos del Espectro-radiómetro de Imágenes de Resolución Moderada (*Moderate Resolution Imaging Spectroradiometer* - MODIS) a los administradores de recursos naturales y a otros actores clave en distintas regiones del mundo. El sensor MODIS se encuentra a bordo de los satélites Terra y Aqua que forman parte de la misión, *Earth Observing System*, de la NASA.

Debido a que los datos asociados a los incendios producidos en la provincia de Córdoba y los datos satelitales vinculados a éstos aún están siendo relevados por el equipo, los resultados y experimentos realizados en el presente trabajo fueron obtenidos mediante la aplicación de las 13 características utilizadas por [Cortez y Morais, 2007] y sobre el área del Parque Natural Montesino de Portugal. En una próxima etapa se procederá a implementar esta nueva información como parte de las características del modelo y su aplicación estará realizada fundamentalmente para la provincia de Córdoba.

Formación de Recursos Humanos

El equipo de investigación y desarrollo de software está formado por docentes investigadores de la Universidad Tecnológica Nacional, Facultad Regional Córdoba. Sin embargo, también participan del proyecto alumnos de los últimos niveles la carrera de Ingeniería en Sistemas de Información de la misma universidad, los cuales realizan sus

prácticas supervisadas (PS, parte de los requisitos para la obtención del grado de Ingeniero) en el marco del proyecto.

Además, año tras año se capacita y forma a alumnos becarios que participan y aprenden desarrollando diversas tareas en el proyecto de investigación, lo que permite complementar su formación curricular desde el punto de vista científico.

Los alumnos intervienen, aportan su trabajo en el proyecto, y aprenden a realizar actividades de investigación, y cómo integrarse en un equipo existente. De esta manera, se está cumplimentando uno de los objetivos del proyecto que es el contribuir a la formación de alumnos en tareas de investigación, en especial fomentando el interés por los sistemas basados en IA.

Asimismo, se plantea como un objetivo adicional del proyecto, que sirva como marco de trabajo para que docentes-investigadores realicen su tesis de Maestría en Sistemas de Información en el presente proyecto.

En síntesis, entre los integrantes que han formado parte del equipo de investigación, se puede precisar:

- 5 docentes investigadores (de los cuales uno realiza su tesis de maestría)
- 2 alumnos auxiliares.
- 4 alumnos becarios.
- 5 alumnos practicantes.

Referencias

[Chandler et al., 1983] Chandler C., P. Cheney, P. Thomas, L. Trabaud, y D. Williams. *Fire in Forestry*, Vol I. *Forest Fire Behaviour and Effects*, John Wiley, New York, p. 450. 1983.

[Cortez y Morais, 2007] Cortez P. y A. Morais. *A Data Mining Approach to Predict Forest Fires using Meteorological Data*. *New Trends in Artificial*

Intelligence, Proceedings of the 13th EPIA 2007 - Portuguese Conference on Artificial Intelligence, Diciembre, pp. 512-523. 2007.

[Cortez, 2010] Cortez P.. Data Mining with Neural Networks and Support Vector Machines Using the R/rminer Tool. In P. Perner (Ed.), Advances in Data Mining - Applications and Theoretical Aspects 10th Industrial Conference on Data Mining (ICDM 2010), pp. 572-583, Springer. 2010.

[Hsu et. Al, 2010] Hsu C.W., C.C. Chang , C.J. Lin y otros. A practical guide to support vector classification. Technical report, 1, Department of Computer Science National Taiwan University, 2010.

[Justicea y et. al., 2002] Justicea C., L. Gigliob, S. Korontzia, J. Owensa, J.T. Morisettec, D. Roya, J. Descloitresb, S. Alleaumed, F. Petitcoline, Y. Kaufman. The MODIS fire products.2002.

[Kim, 2009] Kim H.. Prediction of Forest Fires using Data Mining Methods. M.Sc. Project, University of Western Ontario. 2009.

[Ollero et. al., 2001] Ollero A., J. R. Martínez-de Dios y B.C. Arrúe. Monitorización y determinación de parámetros en tiempo real de fuegos forestales empleando cámaras visuales y de infrarrojos. Actas del III Congreso

Forestal Español. Pp. 503 - 509. Granada.2001.

[Rauste, 1996] Rauste, Y. "Forest Fire Detection with Satellites for Forest Fire Control," Int'l Archives of Photogrammetry and Remote Sensing, Vol.31, Part B7, Proc. XVIII Congress of the Int'l Soc. for Photogrammetry and Remote Sensing, ISPRS, pp. 584-588, 1996.

[Rumelhart, 1986] Rumelhart D. Parallel Distributed Processing: Foundations. MIT Press, 1986.

[Safi y Bouroumi, 2013] Safi Y. y A. Bouroumi. Prediction of Forest Fires Using Artificial Neural Networks. Applied Mathematical Sciences, Vol. 7, Nro. 6, pp. 271 – 286. 2013.

[Van Wagner y Pickett, 1985] Van Wagner y Pickett, Equations and FORTRAN program for the Canadian Forest Fire Weather Index System. Canadian Forestry Service. ISBN 0-662-13906-2. 1985.

[Vega Garcia et. al., 1999] Vega Garcia C., P. Woodard, S. Titus, W. Adamowicz, B. Lee. Dos modelos para la predicción de incendios forestales en Whitecourt Forest, Canadá. Investigación Agraria: Sistemas y Recursos Forestales. pp. 5-24. 1999.