

Modelo predictivo de control en fundiciones de alta precisión: un nuevo enfoque para la fase de predicción^(*)

J. Nieves*, I. Santos* y P.G. Bringas*

Resumen

Un Modelo Predictivo de Control (MPC) es un sistema que permite llevar a cabo el control de una planta de producción. Gracias a este tipo de sistemas es posible realizar una producción que se aproxime al “cero defectos”. Para poder desempeñar su labor fundamental, este tipo de sistemas se compone de varias fases o etapas. Una de las más importantes es la fase que permite predecir la situación en la que se va a encontrar la planta en un momento dado. La mayoría de los avances realizados en este campo han estado ligados a los MPC lineales, a pesar de que el proceso que se intenta representar no lo sea. Así, en este artículo, se presentan varios experimentos que permiten demostrar que la etapa de predicción, habitualmente representada por una única función matemática, puede ser representada mediante modelos de aprendizaje automático.

Palabras clave

Modelo predictivo de control; Aprendizaje automático; Predicción de defectos; Minería de datos; Optimización de procesos.

Model predictive control on high precision foundries: a new approach for the prediction phase

Abstract

A Model Predictive Control (MPC) is a system which allows us to control a production plant. Thanks to this type of system it is possible to make a production that comes close to “zero defects”. In order to achieve its main goal, this kind of systems consists of several phases. One of the most important is the phase that predicts the situation in which the plant is going to be in a given time. Currently, the majority of the research in this field are related to linear MPC, although the process, which the model tries to represent, may not be. Thus, this paper presents several experiments that prove that the forecast phase, usually represented by a single mathematical function, can be represented by machine-learning models.

Keywords

Model predictive control; Machine learning; Fault prediction; Data mining; Process optimisation.

1. INTRODUCCIÓN

Se tiene conocimiento que desde épocas ancestrales existen los procesos productivos. Procesos que han quedado representados tanto en versos bíblicos y pinturas egipcias, como en los grabados de las vasijas griegas. A lo largo de todo este tiempo, los procesos productivos han permitido transformar las materias primas en productos elaborados. Antiguamente, el proceso era mucho más primitivo, sin embargo, en la actualidad, y debido al avance y evolución de la sociedad, el proceso se ha mejorado. Los cambios más

relevantes que se han producido a lo largo de las últimas décadas son: (i) la competencia ha aumentado drásticamente y (ii) las limitaciones impuestas por la legislación medioambiental han dotado de una complejidad extra al proceso de producción.

Estos cambios implican que se lleve a cabo un control estricto de los procesos de producción. De esta forma, y gracias a la ayuda de múltiples disciplinas del conocimiento, es posible anticiparse y solucionar las situaciones adversas del proceso productivo. Así, los aspectos fundamentales sobre los que se puede trabajar son los siguientes: (i) minimizar los costes, (ii) maximizar la tasa de producción, (iii)

^(*) Trabajo recibido el día 2 de Noviembre de 2010 y aceptado en su forma final el día 12 de Abril de 2011.

* Universidad de Deusto, DeustoTech Computing-S3Lab, Avenida de las Universidades 24, 48007 - Bilbao, España.
E-mail: {jnieves, isantos, pablo.garcia.bringas}@deusto.es

minimizar el stock, (iv) controlar la fluctuación de los precios y (v) evitar fallos en el servicio del producto debidos a defectos ocultos.

Esta situación también está afectando al proceso de producción de piezas de metal o proceso metalúrgico. Esta industria es una parte importante de la evolución de nuestra sociedad. Ya desde el punto de vista histórico se ha relacionado el proceso de producción metalúrgico con el desarrollo de la civilización humana, considerándolo incluso como una de las profesiones más antiguas de la historia. La fundición provee a otras industrias, como a la industria automovilística o naval, entre otras. Debido a la naturaleza de estas industrias, hay que realizar grandes controles de seguridad que aseguren el funcionamiento final de las piezas producidas, ya que el menor de los defectos podría ser fatal.

Habitualmente, las medidas para asegurar que el proceso de fundición esté libre de fallos son los controles exhaustivos de la producción y diversas técnicas de simulación. Desafortunadamente, muchas de ellas únicamente pueden ser aplicadas cuando las piezas ya han sido finalizadas. Es por ello, que si se detecta una pieza defectuosa, ésta tendrá que ser fabricada de nuevo, con el incremento de costes del proceso que supone.

Por la problemática presentada, se intenta que los métodos utilizados para determinar el resultado final de una pieza puedan hacerlo incluso antes de fabricarla. En este campo entran los Modelos Predictivos de Control (MPC a partir de ahora). Bajo esta denominación se agrupan una serie de algoritmos capaces de llevar el control de un modelo de proceso concreto. Este tipo de algoritmos trata de mantener la producción de la planta sobre una trayectoria ideal, alineando los parámetros de la producción con la intención de satisfacer las condiciones para las que fue diseñado originariamente el sistema. Sin embargo, a lo largo del proceso pueden darse ciertos cambios que hagan que el resultado final no sea el adecuado. Para solucionarlo, estos algoritmos son capaces de adelantar en el instante de tiempo t cómo se encontrará la planta en el instante de tiempo $t+1$ de seguir así. En caso de detectar algún problema, el sistema se encargará de realizar la corrección de los parámetros manteniendo la producción a lo largo de esa trayectoria óptima. Originariamente, este tipo de sistemas fueron desarrollados para satisfacer las necesidades específicas del control de centrales eléctricas y refinerías de petróleo. Sin embargo, en la actualidad, la tecnología MPC se utiliza en una amplia variedad de áreas como en la producción química, el procesamiento de alimentos, la automoción y las aplicaciones aeroespaciales.

Es fácil confundir los MPC con los sistemas inteligentes para el control de procesos. No obstante, cabe destacar que mientras los segundos disponen de un

conocimiento que les permite tomar decisiones y discurrir acciones según los datos que se están obteniendo de la entrada, los primeros disponen únicamente de conocimiento específico que les permite determinar el estado futuro de la planta mediante métodos de predicción estadísticos. También, mientras que los segundos son capaces de llevar a cabo acciones sobre el sistema debido a la base de conocimiento con la que cuentan, los primeros también llevan a cabo acciones de modificación de parámetros pero basadas en un trabajo de predicciones iterativas. En otras palabras, cuando se ha detectado una desviación de los resultados, se realizan una serie de cambios sobre los parámetros y se vuelve a evaluar el resultado de la planta hasta que la predicción indique que la producción se mantendrá dentro de la trayectoria óptima. En definitiva, ambos sistemas persiguen el mismo objetivo pero para poder conseguirlo llevan a cabo operaciones diferentes. Desde nuestro punto de vista, los MPC son superiores a los sistemas inteligentes de control de proceso debido a que la actualización de los métodos predictivos (o métodos de clasificación) es mucho más sencilla que la del conocimiento experto y concreto del proceso albergado en los sistemas inteligentes para el control de procesos.

En este artículo se presentan varios experimentos de predicción de diversos defectos en el proceso de fundición con el fin de mejorar las etapas de predicción que usualmente se incorporan en los MPC. Además, como objetivos secundarios, se pretende tanto llegar a producir piezas que no fallezcan durante su utilización, así como, reducir el coste del propio proceso de producción al no tener que volver a fabricar las piezas detectadas como defectuosas. Es por ello, que la solución adoptada se traduce en un MPC que deja de lado las formulas matemáticas específicas y que sea capaz de realizar las mismas operaciones a través de modelos de clasificación basados en aprendizaje automático. Particularmente, la comunidad científica se encuentra investigando en la utilización de este tipo de clasificadores con la intención de determinar y prever diversas problemáticas del proceso productivo^[1-4].

En concreto, los experimentos aquí presentados se centran en la predicción de dos de las características más difíciles de prever en las piezas de fundición: (i) los microrrechupes y (ii) las propiedades mecánicas.

Por una parte, el microrrechupe es un defecto o irregularidad en las piezas de fundición que suele aparecer durante la fase de enfriamiento de la pieza. La mayor problemática de éste es que no puede ser detectado hasta que la producción de la pieza haya finalizado. Particularmente, este defecto consiste en una forma de rechupe en el que las cavidades son muy pequeñas pero muchas en número y que pueden estar distribuidas sobre un área importante de la pieza. La

razón de la aparición de este defecto es que los metales son menos densos como líquidos que como sólidos y durante el proceso de enfriamiento, la densidad del metal crece mientras el volumen decrece en paralelo, pudiendo aparecer estas porosidades interdendríticas que reducen la dureza de la pieza haciéndola no apta para su uso. Los procesos actuales para determinar si una pieza tiene este tipo de defecto se basan en la utilización de rayos X o la emisión de ultrasonidos. Ambos métodos necesitan personal especializado, maquinaria adecuada y bastante tiempo para poder realizar los análisis. Por consiguiente, este tipo de soluciones no puede ser considerado como la solución óptima para la producción de piezas sin este tipo de defecto, cuando la alternativa de una detección previa puede mejorar el proceso productivo.

Por otra parte, una vez que las piezas se han producido, éstas tienen que estar sujetas a una serie de fuerzas o cargas. Los ingenieros especializados deben realizar las pertinentes medidas para determinar cómo va a comportarse la pieza. Es decir, si la pieza se va a deformar o romper durante su utilización debido a las fuerzas que se le apliquen. Estas características que cumplen las piezas son las que se conocen como propiedades mecánicas y es importante saber cómo afectan a las piezas de fundición^[5]. Las propiedades mecánicas se pueden agrupar en las siguientes categorías^[6]: (i) fuerza, (ii) dureza, (iii) resistencia, (iv) resiliencia, (v) elasticidad, (vi) plasticidad, (vii) fragilidad, (viii) ductilidad y (ix) maleabilidad. De todas las propiedades mecánicas, en este trabajo nos hemos centrado en la carga máxima. Esta propiedad mecánica pertenece a la categoría de fuerza y es la propiedad que determina la fuerza máxima necesaria para romper la pieza ante una fuerza de tensión sobre la misma. A pesar de que en los experimentos aquí presentados únicamente se hace referencia a una de las propiedades mecánicas, el método utilizado puede ser fácilmente extrapolable para las demás. Al igual que para el defecto anteriormente explicado, las pruebas de control que actualmente se realizan para determinar los valores de las propiedades mecánicas no pueden llevarse a cabo hasta que la pieza está totalmente finalizada. Incluso, en el caso de la carga máxima, estas pruebas acaban con la destrucción de algunas piezas, lo que incorpora nuevos costes al proceso productivo. Por ello, encontrar el enfoque que permita determinar los valores de las propiedades mecánicas antes de realizar las piezas mejorará el proceso productivo.

A la luz de estos hechos, en este artículo se muestra la posibilidad de modelar el proceso productivo mediante una nube de conocimiento adquirido a través de un proceso de aprendizaje automático para posteriormente identificar cuáles de estas aproximaciones pueden ser incluidas dentro de los desarrollos

MPC para mejorar las etapas de predicción. Además, en el caso de utilizar un único algoritmo para la clasificación de los defectos se identifica cuál es el mejor de ellos para cada una de las problemáticas en las que se centra la investigación.

El resto del artículo se encuentra organizado de la siguiente forma. La sección 2 ilustra los modelos predictivos de Control, comentando las limitaciones y lo que se espera en un futuro de ellos. La sección 3 define el proceso de experimentación que se ha llevado a cabo para realizar la mejora de los procesos de predicción de los Modelos Predictivos de Control. Por último, la sección 4 concluye el trabajo y plantea las bases para los futuros desarrollos.

2. MODELOS PREDICTIVOS DE CONTROL (MPC)

Un Modelo Predictivo de Control (MPC) se define como un algoritmo o conjunto de algoritmos que permite controlar cada una de las etapas del proceso productivo de la planta. Para conseguir su objetivo, un MPC basa su funcionamiento en la predicción de cómo se encontrará la planta en un instante de tiempo determinado localizado en un futuro cercano. Los MPC trabajan de forma iterativa. En otras palabras calculan en cada uno de sus ciclos la situación de la planta, determinando si existe alguna variación sobre la producción marcada como idónea. En caso de encontrarla, se modificarán los valores del proceso y se comprobará de nuevo el estado futuro de la planta. Este proceso se repetirá hasta encontrar los valores adecuados para mantener el proceso controlado.

El éxito de la tecnología MPC como un paradigma de control de procesos se puede atribuir a tres factores importantes^[7]. El primero y principal es la incorporación de un modelo de proceso explícito en el cálculo del control. Esto permite que el controlador, en principio, pueda tratar directamente con todas las características importantes de la dinámica del proceso. En segundo lugar, el algoritmo MPC considera el comportamiento de la planta en un horizonte de futuro temporal. Consecuentemente, los efectos de las perturbaciones se pueden prever y eliminar, permitiendo al controlador de la unidad de la planta centrarse en dirigirla a lo largo de la trayectoria deseada. Por último, el controlador MPC considera las entradas del proceso, el estado y las limitaciones de la producción directamente en el cálculo del control de la planta. Por tanto, es casi improbable violar las restricciones del proceso, dando lugar a un control mucho más estricto. Esta inclusión de las limitaciones es lo que distingue los MPC de otros paradigmas de control de procesos.

Sin embargo, aunque los procesos de fabricación, entre los que se encuentra el proceso de fundición, son intrínsecamente no lineales, la gran mayoría de los modelos teóricos desarrollados, al igual que los que se encuentran en el mercado, se basan en modelos dinámicos lineales. Existen varias razones posibles para esto. Primero, los modelos lineales empíricos se pueden identificar de una manera directa a partir de los datos obtenidos mediante una serie de pruebas del proceso. Segundo, la mayoría de las aplicaciones hasta la fecha han sido utilizadas en procesos de refinería^[8], donde el objetivo es, en gran medida, mantener el proceso en un estado de equilibrio deseado en lugar de moverse rápidamente de un punto de trabajo a otro. En otras palabras, el proceso productivo llevado a cabo en las refinerías se mantiene constante, mientras que en otros entornos, como en el de la fundición, el proceso sufre múltiples cambios debido a la gran cantidad de piezas diferentes que pueden llegar a ser manufacturadas, haciendo que los antiguos MPC no se adapten completamente al comportamiento cambiante de la planta. Finalmente, usando un modelo lineal la consecución del objetivo “cero defectos” puede alcanzarse mediante la programación cuadrática, para la que es sencillo encontrar algoritmos fiables^[9].

Por contra, hay casos en que los efectos no lineales son lo suficientemente importantes como para justificar el uso de la tecnología MPC no lineal (NMPC). Estos últimos, son una variante del original MPC que en lugar de utilizar funciones matemáticas lineales para realizar la predicción del estado de la planta, hacen uso de funciones no lineales para obtener su objetivo. Una pequeña problemática que se identifica es que los sistemas MPC deben encontrar su solución en un número finito de iteraciones, algo no tan sencillo con los modelos no lineales. También, los NMPC eran utilizados para entornos con poca tasa de recogida de información, sin embargo, cada vez son más usados en entornos de altas tasas de captura, por lo que este tipo de sistemas puede ser desplegado en una fundición independientemente de cómo sea el entorno de recogida de datos.

Por ende, en este artículo introducimos la utilización de algoritmos de aprendizaje automático. Estos algoritmos permiten trabajar tanto en una dimensión no lineal como en la lineal y todo únicamente dependiendo del algoritmo que se seleccione. De este modo, el objetivo de esta investigación es comprobar el resultado de los algoritmos de aprendizaje automático como sistema de predicción de los modelos MPC.

2.1. Limitaciones

Muchos de los algoritmos MPC, algoritmos que incluso están implantados en la industria, sufren de limitaciones heredadas de sus tecnologías de base. Entre estos problemas se incluyen^[10 y 11]: (i) limitadas opciones entre las que elegir el modelo, (ii) realimentación ineficiente, (iii) pérdida de la estabilidad ante la variabilidad del proceso y (iv) solución ineficiente de la optimización dinámica de los modelos.

Además, las señales de la planta (datos de entrada para un sistema MPC) requieren una atención especial de ingenieros experimentados durante las fases del diseño y prueba del modelo. También, otro de los problemas es que no existe ninguna herramienta para determinar si los datos recogidos son los adecuados para representar el comportamiento del proceso antes de comenzar con el diseño del MPC.

Asimismo, no se aborda la eficiencia estadística y la coherencia de los algoritmos. También, existe una falta de métodos de validación del modelo que permitan comprobar si éste es el adecuado para el control, o si se ha producido un deterioro en su eficacia después de su puesta en marcha. Tampoco existe ningún enfoque sistemático para la construcción de modelos dinámicos no lineales. Por consiguiente, se necesita una serie de directrices para realizar las pruebas de la planta con el objetivo de construir un modelo no lineal confiable.

Estas limitaciones se pueden solucionar en parte gracias a los modelos de aprendizaje automático.

2.2. MPC, la siguiente generación

Según la visión de los vendedores de tecnologías MPC en conjunción con las conclusiones extraídas por Froisy^[12] las áreas de investigación en las que los MPCs deben evolucionar son:

- **Formulación básica del controlador:** debido a la dificultad de expresar todos los objetivos de control pertinentes en función de un solo objetivo, la nueva generación de la tecnología MPC utilizará múltiples funciones objetivo. Por ello, en este trabajo dividimos la problemática con la intención de atacar cada uno de los retos por separado, tratando de solucionar cada una de las funciones objetivo por ella sola.

- **MPC adaptativo:** existen fuertes incentivos del mercado para que la comunidad científica desarrolle un controlador MPC que se autoajuste. Uno de los pocos algoritmos adaptativos es el algoritmo *Control Predictivo Generalizado* (GPC, *Generalized Predictive Control*) introducido por Clarke, Mohtadi

y Tuffs^[13]. Debido a la nueva visión que se está dando de estos tipos de modelos, resulta necesario buscar la forma de readaptar los algoritmos de aprendizaje automático.

- **MPC robusto:** los controladores MPC industriales suelen confiar únicamente en la fuerza bruta de la simulación para evaluar los efectos de la planta. La robustez podría reducir el tiempo necesario para ajustar significativamente el tiempo necesario para adecuar y validar los algoritmos. Incluso, se podrían combinar con las estimaciones de incertidumbre.

- **MPC no lineales (NMPC):** viendo las necesidades del mundo en relación a los modelos predictivos de control, la tendencia de desarrollo de este tipo de sistemas debe dirigirse hacia los NMPC. Y a este tipo de desarrollo ayudan los modelos de aprendizaje automático.

3. PROCESO DE EXPERIMENTACIÓN Y RESULTADOS

Para realizar la validación se han recogido datos de una fundición especializada en el desarrollo de componentes seguros y con un alto nivel de precisión. Los resultados obtenidos en el proceso productivo se envían posteriormente a la industria del automóvil. Esta fundición tiene una producción de alrededor de 45.000 toneladas al año. Como ya se ha comentado, estos experimentos se centran exclusivamente en la predicción del defecto de los microrrechupes y de la carga máxima.

Por otra parte, la aceptación o criterio de rechazo de los modelos estudiados se asemeja a la aplicada por las necesidades finales del cliente. En este sentido, para clasificar los especímenes se han definido una serie de niveles de riesgo posibles. La utilización de estos niveles de riesgo es lo que permite determinar si hay que modificar la tendencia que está siguiendo la producción o no es necesario.

Para el primero de los casos de investigación, los defectos de microrrechupe, se han definido cuatro niveles de riesgo. Los niveles son los siguientes:

- **Riesgo 0:** no se prevé microrrechupe.
- **Riesgo 1:** la probabilidad de que exista microrrechupe es ínfima.
- **Riesgo 2:** existe una probabilidad baja de que aparezca algún microrrechupe en la pieza.
- **Riesgo 3:** la aparición de microrrechupe en la pieza es casi segura.

En este caso se trabajan con cuatro niveles en lugar de con dos (si existe o no el microrrechupe) porque la fundición prefiere aplicar medidas diferentes según lo que se está obteniendo. Los riesgos 0 y 1 no suponen ningún problema para el proceso. Sin

embargo, el riesgo 2 comienza a tenerse en cuenta para la realización de modificaciones en el proceso productivo y el riesgo 3 es el que obligatoriamente debe modificar los parámetros de la producción para evitar la aparición de los microrrechupes.

En cuanto a la predicción de la carga máxima, se ha realizado una división más simple. La razón fundamental de la simplificación de los niveles de riesgo es la adaptación al proceso habitual llevado en la fundición, en la que se marca únicamente un umbral de riesgo. Éstos son los siguientes:

- **Riesgo 0:** se sobrepasa el umbral de 370 Mpa.
- **Riesgo 1:** no se sobrepasa el umbral de 370 Mpa.

A continuación, se va a describir la metodología de experimentación que ha sido llevada a cabo para, posteriormente, incidir en ciertas partes de la metodología intentando profundizar en los detalles relacionados.

3.1. Metodología de experimentación

En concreto, con el fin de evaluar adecuadamente la predicción de los microrrechupes y las propiedades mecánicas se define la siguiente metodología:

- **Adquisición de datos:** proceso mediante el cual se va a extraer el conocimiento de la fundición. Estos datos serán fundamentales para que en las etapas posteriores se pueda realizar el proceso de aprendizaje automático.

- **Validación cruzada:** se ha realizado una validación cruzada de k partes^[14] donde $k = 10$. De esta manera, nuestro conjunto de datos se divide 10 veces en 10 partes diferentes. En cada posible división, se seleccionan 9 de esas partes para la fase de aprendizaje (90 % del conjunto de datos total) y sólo una de las partes para la fase de pruebas (10 % del total de datos). Este proceso se repite durante 10 veces en las que se acaban utilizando todas las partes tanto en la fase de aprendizaje como en la de pruebas.

- **Aprender el modelo:** se ha realizado la fase de aprendizaje de cada algoritmo con cada conjunto de datos de aprendizaje, aplicando diferentes parámetros o algoritmos de aprendizaje según lo que permite cada modelo. Concretamente, se han aplicado los siguientes modelos (para más información sobre los modelos véase la sección):

- *Redes bayesianas (BN del vocablo Bayesian Networks):* para el aprendizaje de las redes bayesianas se han utilizado diferentes algoritmos de aprendizaje estructural (aprendizaje que permite descubrir relaciones entre las variables): $K2$ ^[15], *escalar colinas*^[16] y aprendizaje bayesiano ingenuo aumentado mediante un árbol (TAN, del inglés *Tree Augmented*

Naïve)^[17]. Por otra parte, se han realizado experimentos con el clasificador *bayesiano ingenuo*, que aunque es un algoritmo muy sencillo, en el que se asumen que todas las variables son independientes, suele obtener unos resultados aceptables.

– *K vecinos más próximos (KNN, del nombre inglés K-Nearest Neighbours)*: para el clasificador KNN se han realizado experimentos con $k = 1$, $k = 2$, $k = 3$, $k = 4$ y $k = 5$. No se han continuado las pruebas con más valores de k porque el máximo local de precisión para cada uno de los experimentos posibles se encuentra dentro de este rango de valores.

– *Redes neuronales artificiales (ANN, del vocablo Artificial Neural Networks)*: se ha utilizado una red neuronal compuesta por *múltiples capas de perceptrones (MLP, MultiLayer Perceptron)*^[16] con el algoritmo de *retropropagación*, o en inglés *backpropagation*, para su aprendizaje. Se ha seleccionado únicamente este modelo debido a que es uno de los algoritmos más utilizados en experimentos relacionados con clasificación de defectos.

– *Maquinas de soporte vectoriales (SVM, Support Vector Machines)*: Para entrenar los SVM se han utilizado diferentes funciones de núcleo^[18]: *núcleo polinomial*, *núcleo polinomial normalizado*, *núcleo universal Pearson VII* y *núcleo de función de base radial (RBF, de Radial Basis Function)*.

– *Árboles de decisión (DT, del inglés Decision Trees)*: se han llevado a cabo experimentos con *bosques aleatorios*^[19], una agrupación de árboles de decisión generados al azar que se construye utilizando diferentes cantidades de estos árboles (n) para, de este modo, encontrar el máximo local, igual que sucedía con el algoritmo de clasificación KNN. También se ha utilizado el algoritmo C4.5 para generar un árbol de decisión.

• **Pruebas del modelo**: la primera de las medidas sobre las que se ha trabajado es el nivel de precisión alcanzado por el clasificador. El segundo es la evaluación de las tasas de error. Se ha evaluado la tasa de error entre el conjunto de valores predichos X y el conjunto de valores reales Y (ambos con el tamaño del conjunto de datos de pruebas m) mediante la el error medio absoluto (MAE, *Mean Absolute Error*) (se muestra en la ecuación (1))

$$MAE(X, Y) = \sum_{i=1}^m \frac{|X_i - Y_i|}{m} \quad (1)$$

Del mismo modo, se ha utilizado la medida del error raíz cuadrado (RMSE, *Root Mean Square Error*) (se muestra en la ecuación (2))

$$RMSE(X, Y) = \frac{1}{m} \cdot \sqrt{\sum_{i=1}^m (X_i - Y_i)^2} \quad (2)$$

3.2. Proceso de adquisición de datos

En estos experimentos, los clasificadores se han construido con 24 variables obtenidas del proceso de fundición. Se ha trabajado con varias referencias diferentes, es decir, diferentes tipos de piezas. Con el fin de probar la exactitud de las predicciones, se han utilizado los resultados de inspecciones previas (mediante rayos X y ultrasonidos para los microrrechupes e inspecciones destructivas para las propiedades mecánicas) de entre 889 a 951 especímenes (debe tenerse en cuenta que cada referencia puede incluir más de una pieza de fundición). Cabe destacar que en las primeras fases de la experimentación se contaba con más de 300 variables pero los expertos en el dominio nos ayudaron a reducir la cantidad de variables a manejar.

Para facilitar el proceso de adquisición de datos, el proceso de producción ha sido definido en varias fases. En cada una de esas etapas se recuperan los parámetros de la producción seleccionados para las piezas que se están produciendo en ese momento. En concreto, las etapas en las que se ha dividido el proceso completo son las siguientes:

• **Desarrollo de Patrones**: En esta etapa, los moldes (exteriores) y los núcleos (interiores) son producidos en madera, metal o resina para ser usados en la creación de los moldes de arena donde será construida la pieza final.

• **Moldes de arena y creación de núcleos**: A pesar del hecho de que hay otros métodos, el molde de arena es el método más extendido para la creación de piezas metálicas. La arena se mezcla con arcilla, agua u otros aglutinantes químicos. Posteriormente, máquinas especializadas crean las dos mitades del molde y las unen para generar el contenedor donde el metal fundido será introducido.

• **Fusión y colada**: los metales se funden y se mezclan para que posteriormente puedan ser vertidos en los moldes de arena que darán forma a la pieza.

• **Enfriamiento y separación**: la solidificación de las piezas se controla a lo largo de las líneas de refrigeración hasta que ésta llega a su fin. Una vez la pieza ya ha alcanzado una temperatura adecuada y está totalmente sólida, pasa a limpiarse, para así, retirar todos los restos de la arena del molde. La arena retirada pasa a ser tratada para volver a ser utilizada en futuros moldes.

• **Finalización**: Para finalizar el proceso de creación de las piezas, normalmente se llevan a cabo algunas acciones sobre las piezas, por ejemplo, tratamiento térmico y ratificación de defectos de soldaduras, entre otros.

• **Control final**: una vez que las piezas se encuentran terminadas, son probadas, es decir, se comprueba si hay microrrechupes o si los valores de las propiedades mecánicas son adecuados. Una vez que se ha

determinado el resultado de la pieza, éste se adjunta a la información anterior para poder comenzar con las etapas de aprendizaje automático.

La complejidad de la detección de defectos y características de los productos resultantes surge principalmente del elevado número de variables que participan en el proceso de la producción y, por tanto, pueden influir en el diseño final de los productos. De esta manera, tal y como ya se ha comentado con anterioridad, el número total de variables se ha reducido hasta 24, y más concretamente, las variables que se controlan se pueden dividir en **variables relacionadas con el metal y variables relacionadas con el molde**. Más concretamente, son las siguientes:

- **Variables relacionadas con el metal:**

- Composición: tipo de tratamiento, inoculación y cantidades

- Potencial de nucleación y calidad de la mezcla, obtenido mediante el análisis térmico^[20].

- Vertido: duración del vertido y temperatura.

- **Variables relacionadas con el molde:**

- Arena: tipos de aditivos utilizados para la arena, las características específicas de la arena.

- Molde: parámetros de moldeo y máquina utilizada.

Por lo general, la dimensión y la geometría de la pieza fundida desempeñan un papel muy importante y, por tanto, también se incluyen diferentes variables para mantener el control de estas características. Del mismo modo, el sistema tiene en cuenta los parámetros relacionados con la configuración de cada máquina que trabaja en el proceso de fabricación^[21]. Además, se añaden otras variables como la velocidad de enfriamiento y el tratamiento térmico aplicado a la pieza.

3.3. Clasificadores de aprendizaje automático

A continuación, vamos a describir brevemente los diferentes clasificadores que se han utilizado durante el desarrollo de los experimentos. Estos clasificadores han sido empleados tanto para el descubrimiento de los microrrechupes como para la predicción de propiedades mecánicas.

3.3.1. Redes bayesianas

La investigación sobre las relaciones causa-consecuencia fue iniciada por el reverendo Thomas Bayes^[22] y su trabajo principal es conocido como el “teorema de Bayes” en su honor. De acuerdo a su for-

mulación clásica, dados dos eventos A y B , la probabilidad condicionada de $P(A/B)$, en la que A se produce si se produce B , se puede obtener si se conoce la probabilidad de que A ocurra, $P(A)$, la probabilidad de que B ocurra, $P(B)$, y la probabilidad condicional de B dado A , $P(B/A)$ (tal y como se muestra en la ecuación (3)).

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \quad (3)$$

Extendiendo este modelo, las redes bayesianas son modelos probabilísticos para el análisis multivariable. Podemos representar una red bayesiana como un grafo dirigido acíclico y la función de distribución de probabilidad asociada a ese grafo^[23]. El modelo gráfico representa el conjunto de las relaciones probabilísticas entre las variables recogidas a la hora de modelar un problema en particular. La función de probabilidad ilustra la fuerza de estas relaciones.

Este tipo de modelo ha sido utilizado para varias actividades, por ejemplo, para el desarrollo de una máquina de aprendizaje basado en datos históricos, para la búsqueda de patrones sobre datos ambiguos o incompletos, para la minería de datos a fin de descubrir las relaciones o inferir las variables no observables utilizando el resto del conjunto^[24].

Esta capacidad de inferencia se adapta a nuestros intereses, pues las redes bayesianas representan un superconjunto semántico de los sistemas expertos basados en el encadenamiento de reglas, tanto hacia adelante como hacia atrás (de hecho, los modelos bayesianos permiten un tercer tipo más de inferencia, que se conoce como explicación o justificación^[23]). Además, una red bayesiana puede crecer ampliando su base de conocimiento con nuevas evidencias, sin reducir su nivel de rendimiento^[23] mientras se adapta al problema y mantiene un procedimiento de actualización.

3.3.2. K vecinos más cercanos

El algoritmo de los K vecinos más cercanos (KNN)^[25] es uno de los algoritmos de aprendizaje automático supervisado más simples. Este algoritmo basa su método de clasificación en la clase a la que corresponden la mayoría de los casos más cercanos a la instancia desconocida que va a ser clasificada.

La fase de entrenamiento de este algoritmo realiza una representación de un conjunto de datos de entrenamiento o instancias $S = (s_1, s_2, \dots, s_m)$ en un espacio n -dimensional donde n es la cantidad de variables para cada instancia (en nuestro caso, las 24

variables adquiridas del proceso productivo de la función).

La fase de clasificación de una instancia desconocida, en otras palabras, determinar la clase a la que pertenece, se realiza midiendo la distancia entre las instancias que ya han sido representadas en el espacio n -dimensional y la instancia desconocida. De esta manera, para calcular la distancia entre dos puntos X e Y se calcula mediante cualquier medida de distancia. En nuestro caso, hemos utilizado la distancia euclidiana (véase la ecuación (4)).

$$\sqrt{\sum_{i=0}^n (X_i - Y_i)^2} \quad (4)$$

Durante el proceso de clasificación se pueden utilizar una cantidad diferente de vecinos para realizar la comprobación de sus tipos. No se puede saber de antemano cuál es la cantidad idónea. Por tanto, hay que conseguir descubrir ese número empíricamente. Normalmente, el número de vecinos óptimos suele tomarse a partir del descubrimiento de un máximo local.

Por último, hay varias formas de elegir la clase de la instancia desconocida, la técnica más habitual consiste en clasificar la instancia desconocida como la clase más común entre los k vecinos más cercanos. Esta técnica es la que hemos utilizado en nuestra validación.

3.3.3. Redes neuronales artificiales

Las redes neuronales artificiales (ANN) son un modelo de aprendizaje automático supervisado que simula el comportamiento de las neuronas del cerebro humano^[26]. Formalmente, una red neuronal consiste en una serie de neuronas interconectadas. La activación de una neurona depende de su conjunto de entradas, donde y_j es la activación de la neurona actual, f_j es la función de activación, $W(j, i)$ es el peso de la neurona y a_j es la activación de la neurona de entrada (así se muestra en la ecuación (5)).

$$y_i = f_i \left(\sum_{j=0}^n W_{j,i} \cdot a_j \right) \quad (5)$$

Concretamente, en nuestro estudio, hemos utilizado un tipo de red neuronal formado por múltiples capas de neuronas (MLP). Un MLP es un modelo de red neuronal artificial compuesto por neuronas simples llamadas perceptrones. Estas redes se organizan en diferentes capas tal y como se muestra a continuación:

- **Capa de entrada:** esta capa tiene el mismo número de neuronas que variables definidas en el modelo. Por consiguiente, las variables del modelo son la entrada a las neuronas de esta capa.

- **Capas ocultas:** se crean estas capas para adaptar el aprendizaje, conectando las neuronas de la entrada con las de la salida.

- **Capa de salida:** en la capa de salida va a haber tantas neuronas como grupos de clasificación. En nuestro caso, los niveles de riesgo para microrrechepe y dos rangos de valores para la carga máxima. Según la excitación de estas neuronas se decidirá el resultado de la clasificación de la evidencia.

Para la creación del modelo se utiliza el algoritmo de retropropagación^[27] que calcula los pesos W_j de la función de activación para cada neurona.

3.3.4. Máquinas de soporte vectorial

Las máquinas de soporte vectorial (SVM)^[28] tratan de encontrar un hiperplano o un conjunto de hiperplanos que divida el espacio n -dimensional de los datos en dos o más regiones. El hiperplano es aquél que maximice la distancia entre los ejemplos o evidencias de ambas clases. Esta separación entre regiones se denomina margen funcional. En general, cuanto mayor sea el margen funcional, menor será el error de generalización del clasificador.

Se denominan vectores de soporte a las instancias de las clases que se encuentren más cerca del hiperplano. Dado que a veces el espacio no puede ser dividido con un hiperplano, se hace uso de una función del núcleo K (o función de *kernel*). Esta función estudia las relaciones entre los datos y crea divisiones más complejas en el espacio que las que se pueden generar a través de los hiperplanos. Aunque el algoritmo resultante es formalmente similar, cada producto vectorial, resultante de la representación cuadrática se sustituye por una función de núcleo no lineal. Los SVM se utilizan ampliamente y han conseguido lograr muy buenos resultados en categorización de contenido textual^[29].

3.3.5. Árboles de decisión

Estos clasificadores^[30] constituyen una herramienta de soporte a la toma de decisiones que es representada como un árbol. También, se utilizan para modelar las decisiones y las posibles consecuencias que pueden surgir una vez hayan sido tomadas. Los nodos que componen los árboles de decisión se pueden clasificar en los siguientes tipos:

- **Nodos de decisión:** son los nodos en los que se toma una decisión basada en los valores que se han obtenido previamente. Es decir, se evalúan las condiciones para poder navegar a través del árbol.

- **Nodos de oportunidad:** estos nodos dan la oportunidad de realizar eventos u operaciones que permitan modificar el entorno y que esos cambios puedan tenerse en cuenta en los próximos nodos de decisión.

- **Nodos finales:** los nodos finales son los que determinan el resultado. En nuestro caso, indican si una pieza pertenece a un rango u otro dependiendo de la problemática a tratar (microrrechupes o propiedades mecánicas).

Por otra parte, existen algoritmos de entrenamiento que se utilizan para el aprendizaje de la estructura relacional de estos árboles a partir de un conjunto de datos etiquetados. En este trabajo hemos utilizado los *bosques aleatorios*. Se trata de una agrupación de árboles (es decir, una combinación de clasificadores débiles) construido a partir de diferentes árboles generados al azar^[31]. Además, también hemos utilizado el algoritmo J48 (la implementación del conocido entorno para el aprendizaje automático Weka^[32] del algoritmo C4.5 desarrollado por Ross Quinlan). Los árboles de decisión pueden ser utilizados para mantener la idea de trabajo lineal que desempeñan los habituales sistemas MPC.

3.4. Resultados

Una vez aplicada la metodología mencionada en la sección 3.1, hemos obtenido los resultados que se muestran a continuación. Para facilitar la legibilidad de los mismos, hemos dividido los resultados obtenidos según el defecto objetivo de cada tipo de predicción.

3.4.1. Microrrechupe

Anteriormente, ya se había experimentado la utilización de clasificadores para la predicción del microrrechupe^[33]. Gracias a ese experimento se puede afirmar que este tipo de aproximaciones mejoran el proceso productivo de la fundición. Sin embargo, hasta donde abarca nuestro conocimiento, no existe ningún enfoque que prediga este problema utilizando otro modelo de aprendizaje automático diferente a las redes bayesianas. Por ello, hemos realizado los experimentos para descubrir si realmente las redes bayesianas eran la solución ideal o existía algún clasificador capaz de superar el rendimiento de éstas.

En relación a estos experimentos, la figura 1 muestra los resultados obtenidos en términos de exactitud de la predicción y la figura 2 muestra las tasas de error (MAE y RMSE) de los clasificadores utilizados.

Como se puede observar, casi todos los algoritmos de clasificación obtuvieron buenos resultados, manteniéndose en un porcentaje de acierto que va

desde el 73,29 % hasta el 94,25 %. El mejor de los clasificadores ha sido uno de los clasificadores lineales, los árboles de decisión y, más concretamente, la variante de *bosques aleatorios*. Además, 12 de los 22 clasificadores superan el 90 % de nivel de precisión y 21 de los 22 superan el 80 % de acierto.

De esta manera, casi todos los algoritmos obtuvieron buenos resultados, sin embargo, tanto los árboles de decisión (utilizando *bosques aleatorios* y J48) como las máquinas de soporte vectorial (utilizando *núcleo polinomial normalizado* y *núcleo polinomial*) y las redes bayesianas ingenuas aumentadas mediante un árbol parecen los clasificadores más adecuados si tenemos en cuenta únicamente la precisión.

Por otra parte, el clasificador *bayesiano ingenuo* se comportó como uno de los peores clasificadores. Hay que tener en cuenta que este clasificador es un tipo de red bayesiana en la que todas las variables de entrada se suponen estadísticamente independientes. Por ello, se salta la dependencia causal que pueda surgir entre las variables, por lo que no puede lograr tan buenos resultados como los demás clasificadores. Y, sorprendentemente, el peor de los clasificadores es el que utiliza las máquinas de soporte vectoriales con *núcleo universal Pearson VII*. Este resultado es sorprendente debido a que otros SVM, utilizando diferentes núcleos, se ajustaron mucho mejor al modelo.

También, el algoritmo de los K vecinos más próximos (KNN), un clasificador no lineal, logró mejores resultados de lo que uno puede pensar de antemano. KNN no tiene una fase de entrenamiento o de aprendizaje en sí mismo (únicamente dispone de una pequeña etapa de procesamiento de datos), sólo se centra en la semejanza entre las instancias. Se comporta relativamente bien, aunque los métodos con un proceso de aprendizaje más robusto y que pueden extraer mucha más información de los datos se comportan mucho mejor que este clasificador.

En lo que respecta a las tasas de error, se puede observar que para el MAE (Fig. 2) no se mantiene el mismo comportamiento de los clasificadores que en el nivel de precisión. Mientras que las máquinas de soporte vectorial eclipsan al resto de los clasificadores (con excepción del que utiliza el *núcleo universal Pearson VII*), los árboles de decisión generados mediante *bosques aleatorios* son los que peores resultados obtienen. A pesar de todo, la variación entre los valores obtenidos por los primeros y los segundos es de 0,08 unidades, un valor relativamente pequeño. En el resto de los casos, también se observan variaciones pero que se mantienen en el rango ya comentado. En definitiva, en lo referente a esta tasa de error, todos los clasificadores obtienen resultados parejos.

Por otra parte, y teniendo en cuenta la tasa de error RMSE (Fig. 2), se vuelve a ver que los árboles de deci-

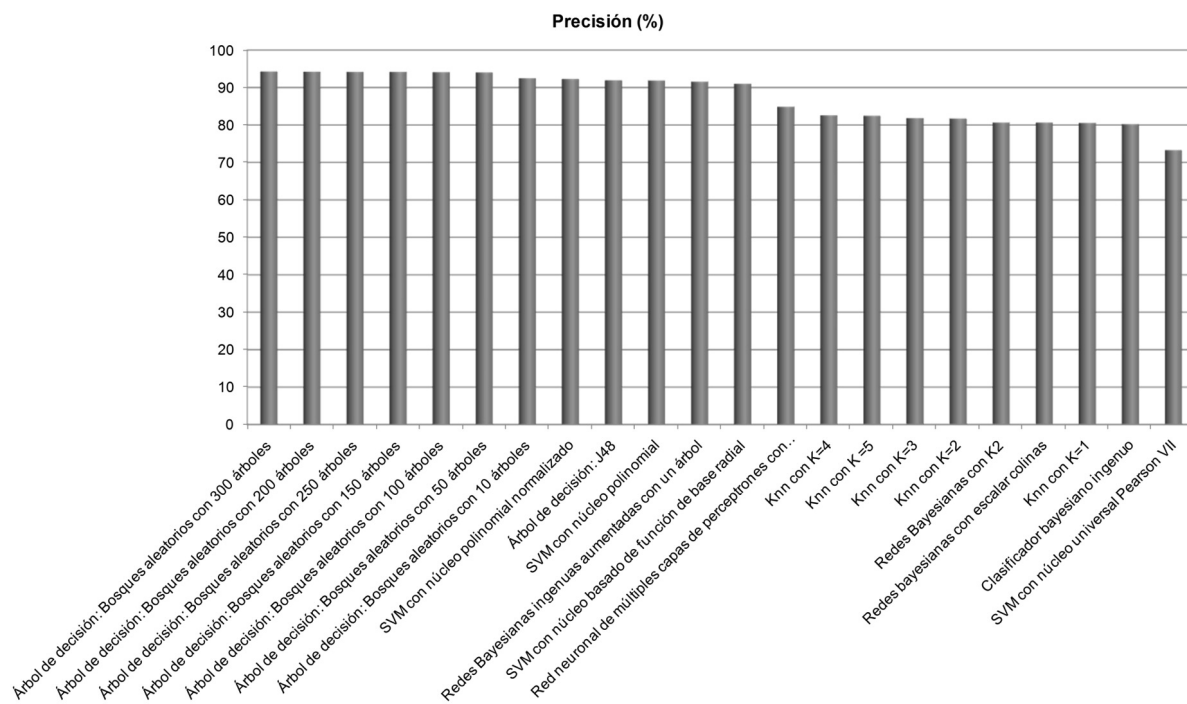


Figura 1. Resultados en términos de nivel de precisión alcanzados por los diferentes clasificadores a la hora de realizar la predicción de los microrrechupes.

Figure 1. Results in terms of accuracy level achieved by several classifiers when making the prediction of microshrinkages.

sión generados con *bosques aleatorios* son los clasificadores que han alcanzado mejores resultados (al igual que en el nivel de precisión). Por eso, estos clasificadores son considerados como los mejores para la predicción de microrrechupes ya que son los más equilibrados. En el caso que nos ocupa, la variación entre todos los clasificadores es mínima ya que la mayoría ronda el rango de valores que van desde 0,23 a 0,26, lo que hace todos ellos tengan un comportamiento similar a las tasas de errores. Sin embargo, y, al igual que en todos los casos anteriores, las máquinas de soporte vectorial con *núcleo universal Pearson VII* son los que peores resultados obtienen. Por lo que este clasificador se podría descartar para la predicción de los microrrechupes como clasificador único. Por contra, y pensando en la posible generación de un meta-clasificador, o conjunto de clasificadores que trabajan unidos para alcanzar el objetivo común, sí es interesante que forme parte de la comunidad y que se tenga en cuenta su opinión para intentar realizar la predicción de resultados.

En resumen, este tipo de clasificación basada en técnicas de aprendizaje automático (que en la mayoría de los casos ha alcanzado más del 90 % de precisión) es una interesante solución a adoptar para la predicción de los microrrechupes. De esta manera, se reduce de manera significativa el coste y la duración

de los métodos de las pruebas actuales, así como se asegura la reducción del número de piezas que pueden desfallecer mientras se está llevando a cabo su uso. Cabe destacar que los *bosques aleatorios* pueden ser usados de la misma forma en la que se usaban las redes bayesianas en los primeros de los experimentos^[33]. Además, nótese que aunque se buscan nuevos modelos de MPC no lineales, el clasificador que mejor resultado ha obtenido es inherentemente lineal. La razón fundamental es que la generación de *bosques aleatorios* no es la habitual en la generación de un árbol de decisión. Los *bosques aleatorios* son una agrupación de clasificadores que por sí solos se sabe que van a funcionar con peor rendimiento, pero que al trabajar en grupo obtienen mejores resultados. A pesar de todo, la variación entre estos clasificadores lineales y los no lineales con mejores resultados es únicamente de 4 unidades. Por tanto, la combinación de clasificadores, como hacen los *bosques aleatorios*, puede ser una la mejor solución para la predicción de microrrechupes.

3.4.2. Propiedades Mecánicas

Sobre las propiedades mecánicas, en relación a nuestro conocimiento, no existe ningún modelo que per-

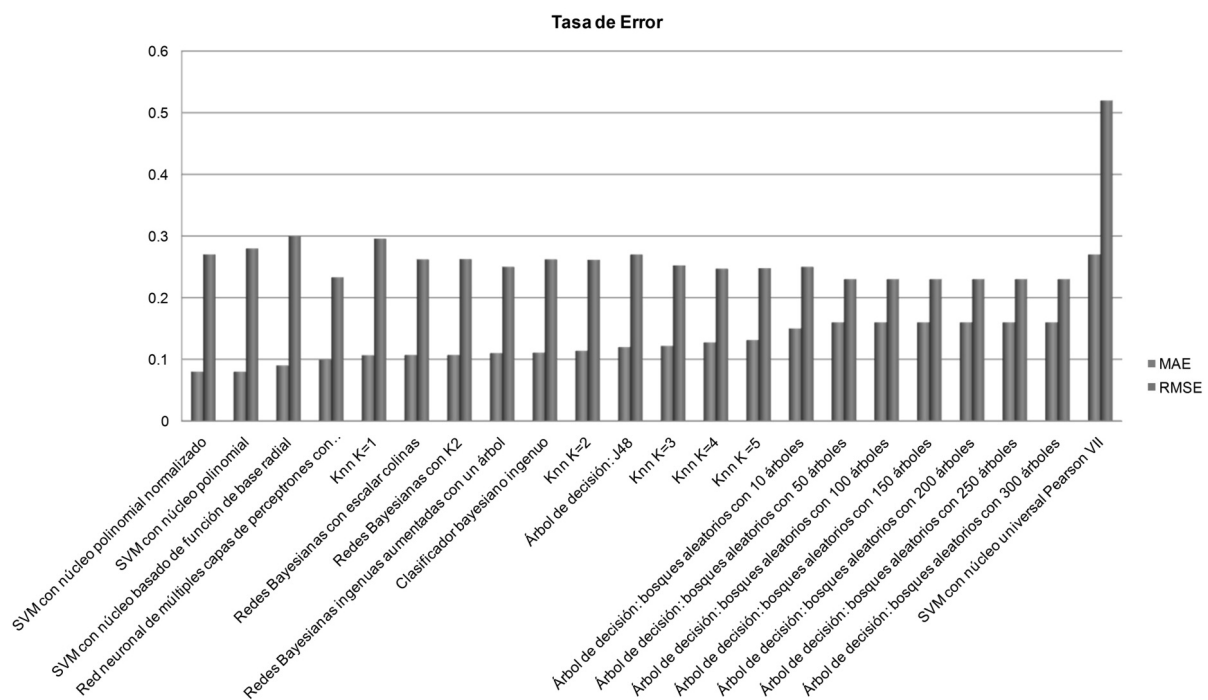


Figura 2. Resultados en términos de tasa de error alcanzados por los diferentes clasificadores a la hora de realizar la predicción de los microrrechupes, donde MAE es la tasa de error medio absoluto y RMSE es la tasa de error raíz cuadrado.

Figure 2. Results in terms of error rates achieved by several classifiers when making the prediction of microshrinkages, in which MAE is the Mean Absolute Error and RMSE is the Root Mean Square Error.

mita predecir su estado antes de llegar a realizar las piezas. Por eso, se decidió dar los mismos pasos que con los microrrechupes.

Uno de los aspectos que merecen ser comentados es la existencia de un modelo previo de redes bayesianas, así que, el primero de los pasos fue replicar el experimento con el fin de comparar los resultados. Es por ello que en la figura 3 y 4, se puede ver un clasificador más que en las de los microrrechupes. En este caso, el clasificador que ha sido añadido es la red bayesiana en la que para realizar el aprendizaje estructural se utiliza el algoritmo PC o el NPC. En particular, el algoritmo PC [34] se utiliza aquí para conseguir la estructura de causalidad y/o relaciones correlativas entre las variables que se reciben de los datos de la fundición. En otras palabras, el algoritmo PC utiliza los datos de la muestra de entrenamiento para definir el modelo bayesiano, lo que representa el conjunto de la dependencia y las relaciones de independencia entre los parámetros. Si se conoce de antemano alguna de esas relaciones entre las variables y éstas deben aparecer en el modelo bayesiano, podemos aplicar el algoritmo NPC [35]. El algoritmo NPC permite definir estas relaciones iniciales que son conocidas como *condiciones de ruta necesarias*.

En el caso de la carga máxima, los árboles de decisión generados con *bosques aleatorios* obtuvieron los mejores resultados en términos de precisión. Para la predicción de la carga máxima se consiguió alcanzar un nivel de precisión del 86,84 % utilizando 250 árboles. Este es el número de árboles seleccionado como óptimo ya que como ilustra la figura 3 se trata del máximo local. A pesar de ser el mismo clasificador que obtuvo los mejores resultados para los microrrechupes, esta vez fue generado con menos árboles. Por lo que no podemos utilizar los mismos parámetros de configuración para ese clasificador. Por consiguiente, para cada uno de los problemas hay que separar el modelo experto que se encarga de la predicción de los defectos.

Muy de cerca siguen las redes neuronales artificiales que obtuvieron un buen nivel de precisión (un 84,23 %). La diferencia es mínima, únicamente se reduce el nivel de precisión en 2,61 unidades. Con lo que podemos considerar que también se obtienen buenos resultados mediante clasificadores no lineales.

Es sorprendente el comportamiento que tuvo el algoritmo de los K vecinos más cercanos. A pesar de su simplicidad, los valores de $k = 1$ y $k = 5$ alcanzaron unos resultados muy similares a los de los árboles de decisión y redes neuronales artificiales. Sin embargo, el resto de

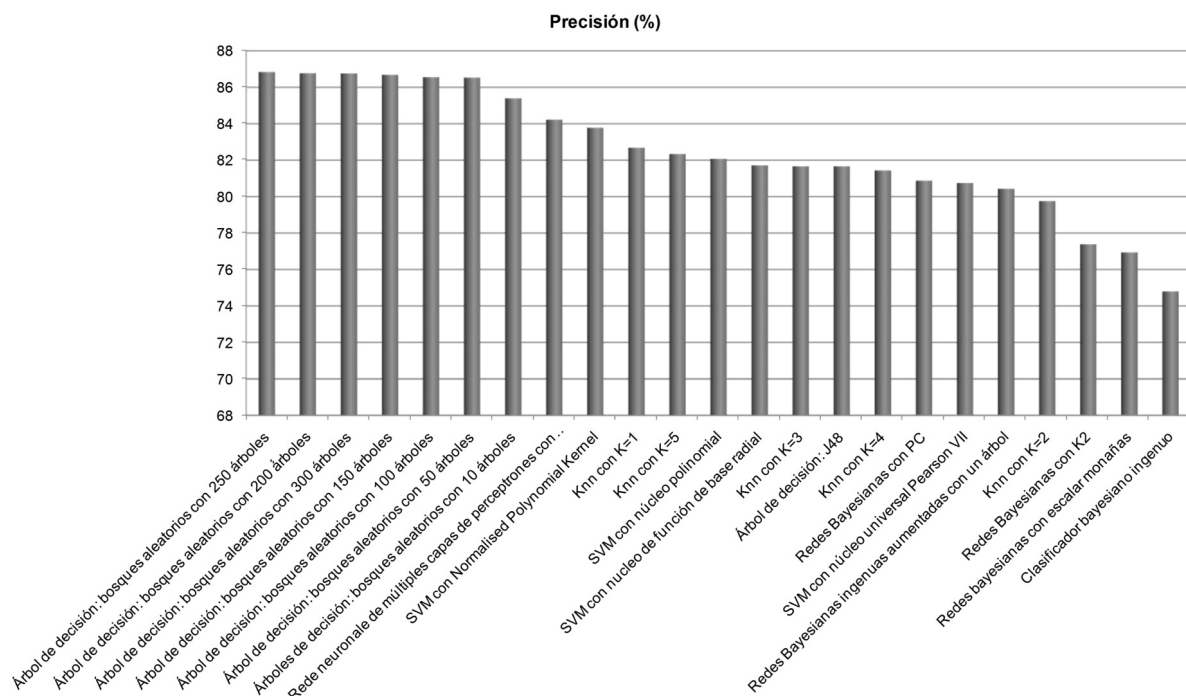


Figura 3. Resultados en términos de precisión alcanzados por los diferentes clasificadores a la hora de realizar la predicción de la carga máxima, un tipo de propiedad mecánica.

Figure 3. Results in terms of accuracy achieved by several classifiers when predicting the ultimate tensile strength, a type of mechanical properties.

los valores de K no lograron unos valores tan interesantes. Por lo tanto, este modelo es una buena opción a tener en cuenta a la hora de desarrollar un meta-clasificador. No obstante, en caso de querer trabajar con un único clasificador para la predicción de las propiedades mecánicas no sería la opción más adecuada.

En el caso de la predicción de la carga máxima, cabe destacar el mal resultado que lograron los modelos basados en clasificadores bayesianos. La mayoría de las aproximaciones basadas en este tipo se encuentran entre los clasificadores con peor comportamiento. Destaca que a pesar de que la naturaleza del problema es similar y que muchas de las variables con las que se trabajan son las mismas que para la predicción de los microrrechupes, no son capaces de alcanzar el mismo nivel de resultados. Aspecto harto interesante debido a que nos indica claramente que no se puede presuponer que un mismo clasificador se comporte de forma similar a pesar de que la gran mayoría del conocimiento con el que se ha generado sea similar.

En lo referente a las máquinas de soporte vectorial, su comportamiento parece desigual, ya que unos se postulan como la mejor alternativa para los árboles de decisión y redes neuronales artificiales (aquéllos que son generados con el *núcleo polinomial normalizado*), mientras que otros no son capaces de superar clasifi-

cadore tan simples como el algoritmo de los K vecinos más cercanos. También, al igual que para los microrrechupes, el peor de los clasificadores basados en máquinas de soporte vectorial es el que se creó utilizando el *núcleo universal Pearson VII*, lo que hace pensar que no es capaz de adaptarse a los datos que son extraídos de la fundición. De todas formas, al igual que ya se comentó para la predicción de los microrrechupes, es interesante la utilización de este clasificador en el conjunto de clasificadores que formen parte de un meta-clasificador ya que puede compartir con el resto de los clasificadores el conocimiento alcanzado.

En cuanto a las tasas de error, para el MAE (Fig. 4) nos encontramos en la misma situación que en la predicción de los microrrechupes. Ninguno de los clasificadores mostró el mismo comportamiento alcanzado en la precisión. A pesar de todo, los valores de las tasas de error van desde 0,16 unidades a 0,27 (una variación de 0,11 unidades) siendo de poca importancia la diferencia del comportamiento que lograron cada uno ellos. Sin embargo, hablando de la tasa de error RMSE (Fig. 4), los resultados obtenidos se acercan más a los de precisión. Al igual que en el caso del MAE, la diferencia de los valores es muy baja (los valores están en el rango que va de las 0,31 unidades a las 0,44 unidades). Pero en este caso, los árboles de

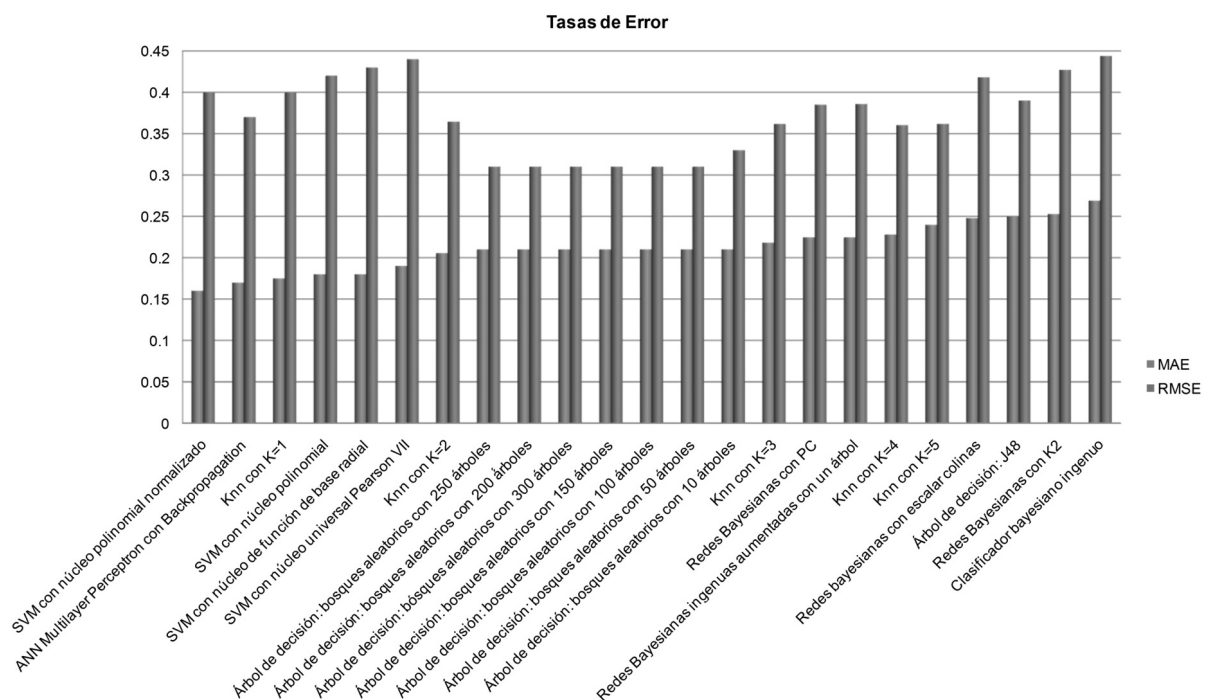


Figura 4. Resultados en términos de tasa de error alcanzados por los diferentes clasificadores a la hora de realizar la predicción de la carga máxima (una de las propiedad mecánica), donde MAE es la tasa de error medio absoluto y RMSE es la tasa de error raíz cuadrado.

Figure 4. Results in terms of error rates achieved by several classifiers when making the prediction of ultimate tensile strength (one of the mechanical properties), in which MAE is the Mean Absolute Error and RMSE is the Root Mean Square Error.

decisión vuelven a ser los mejor clasificadores. Así, los podemos considerar como el método más equilibrado.

Como conclusión, aunque la predicción de la carga máxima no alcanza el mismo nivel de precisión que en la predicción de microrrechupes, se puede considerar que este tipo de solución es adecuada, ya que con el nivel alcanzado (un 86,84 %) los beneficios que puede aportar al sistema de producción de la fundición son elevados. También, se ha detectado que existen problemas en los datos extraídos de la fundición, fallos introducidos por los operarios que se encargaban de realizar las anotaciones. Este ruido en la muestra de entrenamiento hace que el comportamiento de los clasificadores no sea el más adecuado. Esto es así, porque al no estar trabajando con todas las medidas tomadas correctamente se está distorsionando la realidad de los clasificadores. Una vez que se haya solucionado este problema, los resultados obtenidos serán mejores debido a que la representación será más fiel al modelo real de la fundición.

En el caso de tener que utilizar un clasificador para la predicción de este tipo de características, lo más idóneo sería la utilización de los árboles de decisión generados con *bosques aleatorios*, ya que con los experimentos aquí realizados se demuestra que

este tipo de aproximación obtiene buenos resultados y que podría ser utilizado para sustituir los actuales sistemas de predicción utilizados en los MPC.

5. CONCLUSIONES

- En este artículo hemos incluido y adaptado a nuestro dominio de aplicación particular diferentes clasificadores que han sido ampliamente utilizados en cuestiones similares. Todos ellos se comportan bien, pero los *árboles aleatorios* superan al resto de los clasificadores. Viendo los resultados obtenidos se puede indicar que la utilización del aprendizaje automático es una opción viable para la mejora de los procesos de predicción de los MPC. Con ello, se consigue obtener una serie de modelos MPC no lineales, pero incluso, generando un meta-clasificador, se podría conseguir que el MPC sea híbrido solucionando parte de los problemas del área.
- Nótese que a lo largo de la validación empírica hemos observado que existen algunas irregularidades en los datos que pueden alterar el resultado generadas durante el proceso de adquisición de datos. Esto abre una nueva línea de trabajo futuro para la

eliminación del ruido mediante la aplicación de técnicas como *Análisis de Componentes Principales* (PCA del vocablo *Principal Component Analysis*) o *Análisis Semántico Latente* (LSA de la voz inglesa *Latent Semantic Analysis*). También, existe la necesidad de un proceso de adaptación de estos clasificadores, abriendo así, otra línea de trabajo sobre técnicas que mantengan la información del modelo pero con un menor nivel de importancia y añadan la nueva información recuperada. Este tipo de técnicas aplican factores de envejecimiento para conseguir que la adaptación del conocimiento sea mucho más rápida.

REFERENCIAS

- [1] A. Lazaro, I. Serrano, J. P. Oria, y C. de Miguel. *Proc. AMC*. Coimbra, Portugal, IEEE Standards Office, Cambridge, MA, EE.UU., pp. 653-658.
- [2] P. Zhang, Z. Xu, y F. Du, *Proc. ICNC*, IEEE Computer Society, Los Alamitos, CA, EE.UU., 2008, pp. 545-548.
- [3] M. Perzyk y A. Kochanzski, *P. I. Mech. Eng. B-J. Eng.* 217 (2003) 1.279-1.284.
- [4] H. Bhadeshia, *ISIJ Int.* 39 (1999) 966-1.000.
- [5] R. Gonzaga-Cinco y J. Fernández-Carrasquilla, *Rev. Metal. Madrid* 42 (2006) 91-102.
- [6] C. W. Lung y N. H. March, *Mechanical Properties of Metals: Atomistic and Fractal Continuum Approaches*, Ed. World Scientific Pub Co Inc, Singapur, 1999, pp. 43-110.
- [7] M. Morari y J. H. Lee, *Comput. Chem. En.* 23 (1997) 667-682.
- [8] S.J. Qin y T.A. Badgwell, *Prog. Syst. C.* 26 (2000) 369-392.
- [9] S.J. Wright, *Proc. CPC-5*, vol. 93, Tahoe, J.C. Kantor, C.E. Garcia, y B. Carnahan (Eds.), American Institute of Chemical Engineers, Nueva York, EE.UU., 1997, pp. 147-155.
- [10] M. Morari y J.H. Lee, *Proc. CPC-4*, 1991, Padre Island, Texas, EE.UU., Y. Arkun, y W. H. Ray (Eds.), Elsevier, Amsterdam, Holanda, 1991, pp. 419-444.
- [11] K.R. Muske y J.B. Rawlings, *AIChE. J.* 2 (1993) 262-287.
- [12] B. Froisy. *ISA T.* 33 (1994) 235-243.
- [13] D. W. Clarke, C. Mohtadi y P. S. Tuffs, *Automatica* 2 (1987) 137-148.
- [14] R. Kohavi. *Int. Joint Conf. Artif.* (1995) 1.137-1.143.
- [15] G. F. Cooper y E. Herskovits, *Proc. UAI*, Los Angeles, B. D. D'Ambrosio, P. Smets, P. P. Bonissone (Eds.) Morgan Kaufman Publishers Inc., San Francisco, EE.UU. 1991, pp. 86-94.
- [16] S. J. Russell y Norvig. *Artificial Intelligence: A Modern Approach*, Ed. Prentice Hall, Upper Saddle River, NJ, EE.UU., 2003, pp. 126-129.
- [17] D. Geiger, M. Goldszmidt, G. Provan, P. Langley y P. Smyth, *Mach. Learn.* 2 (1997) 131-163.
- [18] B. Üstün, W.J. Melssen y L.M.C. Buydens. *Anal. Chim. Acta*, 595 (2007) 299-309.
- [19] L. Breiman. *Mach. Learn.* 45 (2001) 5-32.
- [20] P. Larrañaga, J. Sertucha, y R. Suárez, *Rev. Metal. Madrid* 42 (2006) 244-255.
- [21] J. Sertucha, R. Suárez, J. Legazpi, y P. Gacetabeitia, *Rev. Metal. Madrid* 43 (2007) 188-195.
- [22] T. Bayes, *Philos. Trans. R. Soc.* 53 (1763) 370-418.
- [23] E. Castillo, J. M. Gutiérrez y A. S. Hadi, *Expert Systems and Probabilistic Network Models*. Ed. Springer, New York, EE.UU., 1996, pp. 69-534
- [24] J. Pearl, *Handbook of brain theory and neural networks*, M. Arbib (Ed), MIT Press, Cambridge, MA, EE.UU., pp. 149-153
- [25] E. Fix y J. L. Hodges, Discriminatory analysis: Nonparametric discrimination: Small Sample performance, *Technical Report Project 21-49-004, Report Number 11*, USAF School of Aviation Medicine, Randolph Field, Texas, EE.UU., 1952.
- [26] C. M. Bishop, *Neural Networks for Pattern Recognition*, Ed. Oxford University Press, Oxford, UK, 1995, pp. 225-283.
- [27] D. Michie y D. J. Spiegelhalter, *Machine learning, neural and statistical classification*, Ed. Ellis Horwood, Upper Saddle River, NJ, EE.UU., 1994, pp. 84-124.
- [28] V. N. Vapnik, *The nature of statistical learning theory*, Ed. Springer, New York, EE.UU., 1995, pp. 156-167.
- [29] T. Peng, W. Zuo y F. He, *Knowl. Inf. Sys.* 3 (2008) 281-301.
- [30] J. R. Quinlan, *Mach. Learn.* 1 (1986) 81-106.
- [31] L. Breiman, *Mach. Learn.* 1 (2001) 5-32.
- [32] S.R. Garner, *Proc. de la New Zealand Computer Science Research Students Conference*, Nueva Zelanda, 1995, pp. 57-64.
- [33] Y. K. Peña, P. G. Bringas y A. Zabala, *Proc. INDIN-6*, Daejon, Corea, 2008, pp. 1.673-1.677.
- [34] P. Spirtes, C. Glymour y R. Scheines, *Causation, Prediction, and Search*, Ed. MIT Press, Cambridge, EE.UU., 2001, pp. 73-156.
- [35] U. B. Kjærulff y A. L. Madsen, *Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis. Information Science and Statistics*, Ed. Springer, New York, EE.UU., 2008, pp. 241-245.