

Tipo de artículo: Artículo original  
Temática: Inteligencia Artificial  
Recibido: 20/05/2017 | Aceptado: 25/06/2017

## Evaluación del perfil neurotóxico de líquidos iónicos usando técnicas de aprendizaje supervisado

### *Evaluation of the Neurotoxic Profile of Ionic Liquids Based using Supervised Machine Learning Techniques.*

Maykel Cruz Monteagudo <sup>1</sup>, Rotceh Dominguez López <sup>2</sup>, Ariel Céspedes Pérez <sup>3\*</sup>, Ricardo Enrique Pérez Guzmán <sup>4</sup>

<sup>1</sup> Universidad de Porto, gmaikelcm@yahoo.com, Portugal.

<sup>2</sup> Universidad de Las Tunas, {rdominl2003, arielcespedes87}@gmail.com, {arces, ricardopg}@ult.edu.cu, Cuba.

\* Autor para correspondencia: [arielcespedes87@gmail.com](mailto:arielcespedes87@gmail.com)

---

#### Resumen

La enzima Acetilcolinesterasa (AChE) juega un papel imprescindible en la hidrólisis del neurotransmisor Acetilcolina, el cual es el responsable de la transmisión de los impulsos nerviosos. Desde la década de los 30, especialistas en las ciencias químicas han producido compuestos que son capaces de inhibir esta enzima y por tanto afectar el proceso de transmisión de los impulsos nerviosos, lo cual provoca consecuencias graves para el organismo afectado. Estudios actuales han demostrado que algunos líquidos iónicos pueden inhibir el funcionamiento de la enzima AChE y provocar daños al sistema nervioso central. Los líquidos iónicos debido a sus características físico-químicas son ampliamente utilizados en la producción de solventes que son utilizados en la sustitución de solventes moleculares tóxicos para el medio ambiente. En correspondencia con esto surge la necesidad de evaluar el perfil neurotóxico de los líquidos iónicos utilizando la enzima AChE como indicador de neurotoxicidad. En el desarrollo del trabajo se aplicaron multclasificadores, como técnicas de aprendizaje supervisado, y como resultado se obtuvieron modelos capaces de predecir si un nuevo líquido iónico es capaz de inhibir la AChE. El multclasificador AdaBoostM1, que utiliza una red neuronal MultilayerPerceptron como clasificador base y el multclasificador Stacking, que utiliza la combinación de clasificadores FDLA, Jrip, Kstar, NaiveBayes y SMO como clasificadores bases, fueron los multclasificadores seleccionados.

**Palabras clave:** Acetilcolinesterasa, líquidos iónicos, QSAR, multclasificadores, medidas de diversidad.

### **Abstract**

*The enzyme Acetylcholinesterase (AChE) plays an essential role in the hydrolysis of the neurotransmitter Acetylcholine, which is responsible for the transmission of nerve impulses. Since the 1930s, specialists in the chemical sciences have produced compounds that are able to inhibit this enzyme and therefore affect the transmission process of nerve impulses, which causes serious consequences for the affected organism. Current studies have shown that some ionic liquids can inhibit AChE enzyme function and cause damage to the central nervous system. Ionic liquids due to their physical-chemical characteristics are widely used in the production of solvents that are used in the substitution of molecular toxic solvents for the environment. In correspondence to this arises the need to evaluate the neurotoxic profile of ionic liquids using the AChE enzyme as an indicator of neurotoxicity. In the development of the work multiclassifiers were applied as supervised learning techniques, and as a result models were obtained capable of predicting if a new ionic liquid is able to inhibit AChE. Bagging, Boosting, Stacking and Vote multiclassifiers were used in the experimentation to identify predictive QSAR models. Five measures of diversity were calculated for the base classifiers used in Stacking and Vote multiclassifiers. Finally, two models were obtained that surpassed the performance of the individual classifiers used, reason why they were selected to solve the problem. The multiclassifier AdaBoostM1, which uses a Multilayer Perceptron neural network as the base classifier and the Stacking multiclassifier, which uses the combination of classifiers FDLA, Jrip, Kstar, NaiveBayes and SMO as base classifiers, were the multiclassifiers selected.*

**Keywords:** *Acetylcholinesterase, ionic liquids, QSAR, ensembles, diversity measures*

---

## **Introducción**

La acetilcolina (ACh) es un neurotransmisor en muchas sinapsis colinérgicas del sistema nervioso autónomo (J. Koolman 2005). La acetilcolinesterasa (AChE) es una esterasa que hidroliza a la acetilcolina después de que ésta ha realizado su función en la unión de receptores. La función normal de la ACh depende de su rápida hidrólisis por la AChE, que permite la brevedad y unidad de los impulsos nerviosos propagados sincrónicamente (Mathews. 2005).

La inhibición de la enzima AChE provoca que la acetilcolina unida a los receptores presentes en la membrana postsináptica no sea liberada rápidamente, como consecuencia la acetilcolina sigue siendo producida como parte de la transmisión de los nuevos impulsos nerviosos sin embargo los receptores no son capaces de recibir los nuevos impulsos eléctricos emitidos por las neuronas vecinas (debido a que la AChE no hidroliza los neurotransmisores ACh) provocando efectos adversos tales como: enfermedades cardiovasculares, debilidad muscular y en ocasiones la muerte (Metzler 2003; Stock y cols. 2004).

Los líquidos iónicos (LI) constituyen un área de las ciencias químicas que han estado atrayendo la atención entre los investigadores debido a las características físico-químicas únicas de estos compuestos que proveen un amplio rango de aplicaciones (Störmann y cols. 2014).

El número de posibles combinaciones de aniones y cationes que pudieran componer un LI está por el orden de los  $10^{18}$  (Arning, S. S. J. y cols. 2008). Los LI son catalogados como “verdes” y la razón consiste generalmente en tres argumentos: i) poseen una presión de vapor muy baja que proporciona una reducción de la exposición por inhalación en el personal que trabaja con él, en comparación con disolventes moleculares convencionales; ii) no son inflamables, lo que reduce fuertemente el riesgo de oxidaciones exotérmicas rápidas y iii) se consideran no tóxicos. La mayor parte de los investigadores expresan su desacuerdo con la última característica y la justificación de su discrepancia se encuentra plasmada en varios artículos (Frade y Exp. 2010; García-Lorenzo y cols. 2008; Zhao y Zhang 2007). Un creciente número de estudios, que analizan el peligro de diferentes LI, ha sido realizados y determinados en diferentes sistemas de pruebas biológicas (J. P. J. S. Torrecilla 2010; Sierra; Martí y Cruañas 2006). A partir de los resultados de estas investigaciones se demuestra que existen LI que poseen un peligro potencial para el hombre y para el medioambiente. Dentro del conjunto de sistemas de pruebas realizadas se encuentra el estudio denominado “Ensayos sobre inhibición de la enzima Acetilcolinesterasa”, realizado a partir de la enzima purificada de la manta eléctrica *Electrophorus electricus*, apoyados los estudios QSAR (Störmann y cols. 2014). Esta prueba se realiza teniendo en cuenta que esta enzima es parte esencial del sistema nervioso central de los organismos complejos (Stock y cols. 2004). Los resultados de este estudio permitieron determinar que los LI pueden inhibir la enzima y como consecuencia provocar efectos no deseados. Los resultados de los ensayos in vivo son un indicador de que la AChE de otros organismos también puede ser inhibida (Stock y cols. 2004). Como resultado de las investigaciones sobre el efecto inhibitorio de los LI sobre la AChE se ha generado un gran volumen de registros que describen las características estructurales de los LI ya sintetizados y la actividad biológica que tienen sobre la AChE. Teniendo en cuenta que la mayoría de los LI aún no han sido sintetizados, es necesario desarrollar métodos para predecir el peligro potencial de LI desconocidos con el objetivo de facilitar el diseño de nuevos materiales seguros y reducir la necesidad de realizar las síntesis de LI basados en el método de prueba y error, y de esta forma disminuir los costos y el tiempo de su producción.

La computación, como herramienta, complementa la realización de estudios que, a partir del uso de simulaciones o procesamiento de grandes volúmenes de datos, que de forma manual sería poco factible realizar. El aprendizaje automático es un área de la computación (específicamente de la inteligencia artificial) en la que el hombre ha obtenido buenos resultados aplicada al procesamiento de grandes volúmenes de información y en la actualidad está comenzando

a ser aplicada con éxito en problemas similares en la rama de la química. En este sentido, pocos artículos relacionados con herramientas computacionales y su utilización para la predicción de la toxicidad de los LI han sido reportados en la literatura, (Arning, S. S. J. y cols. 2008; Couling y cols. 2006; García-Lorenzo y cols. 2008; Hossain y cols. 2011; Izadiyan 2011; Lacrama 2007; Luis, A. G. P. y Irabien 2010; Luis, I. O. P.; Aldaco y Irabien 2007; Mohammad y Fatemi 2011; Putz 2007; Torrecilla y cols. 2010) en particular aquellos relacionados con el potencial inhibitorio de los LI sobre la enzima AChE, como un problema de clasificación.

## **Materiales y métodos**

El desarrollo del trabajo se realizó aplicando un enfoque de minería de datos. Se ejecutaron las etapas del proceso de extracción del conocimiento en base de datos, con el objetivo de encontrar modelos que permitan identificar patrones a partir de la información actual y a partir de esta predecir el peligro potencial de LI desconocidos.

### **Colección y estandarización de los datos**

Los datos utilizados fueron obtenidos de la Base de datos de efectos biológicos de LI pertenecientes a la UFT/Merck (Arning, S. S. Jurgen y Boshen 2008; Stock y cols. 2004; Störmann y cols. 2014). En Tropsha y Muratov (2010), Tropsha y colaboradores proponen una metodología para preprocesar la información relacionada con la estructura de moléculas químicas que serán objeto de la realización de estudios QSAR. Proponen eliminar todas aquellas moléculas que poseen elementos metálicos, organometálicos y elementos poco representados, dejando fuera las moléculas que contienen elementos que no son manejados por las herramientas que calculan descriptores moleculares y aquellas que poseen elementos poco representados que pudieran constituir ruidos en la fase de aprendizaje. Los elementos poco representados son determinados con la utilización del software Jchem (Chemaxon 2012) el cual posee funciones para determinar los elementos que están contenidos en una molécula y con esta información determinar el total de elementos que estaban contenidos en las moléculas.

La estructura química de cada molécula, que se encuentra codificada con SMILES1, debe ser estandarizada debido a que en ocasiones un mismo grupo funcional puede ser representado por diferentes patrones estructurales en un conjunto de datos dado. Para los estudios QSAR esta situación conduce hacia serios problemas debido a que los descriptores moleculares calculados a partir de estas representaciones diferentes del mismo grupo funcional pueden ser significativamente distintos (Tropsha y Muratov 2010) de forma que estos se calculan con si fueran moléculas diferentes. La estandarización estructural de las moléculas se puede realizar utilizando la herramienta ChemAxon's Standarizer (Chemaxon. 2012) debido a que permite realizar normalizaciones de forma rápida y eficiente (Tropsha y

Muratov 2010). El resultado de esta operación permitirá identificar si existen duplicados de una molécula y por lo tanto se está en condiciones de realizar la eliminación de las moléculas duplicadas, debido que pueden afectar la frecuencia observada de una molécula y la distribución de los compuestos acorde a su similaridad estructural. Para la eliminación de duplicados se utilizó el software ISIDA's EdiSDF (V. Solov'Ep 2010), debido a que es una herramienta libre y con fines académicos (Tropsha y Muratov 2010).

El cálculo de los descriptores moleculares se realizó utilizando la herramienta ISIDA (Isida 2008), permitiendo la obtención de un vector numérico que representa la ocurrencia de fragmentos presentes en la molécula. Posibilitando la obtención de información que podrá ser utilizadas por la herramienta de minería de datos a utilizar.

### **Diseño del experimento**

El conjunto de datos resultante recoge la información de 261 moléculas, donde cada molécula está representada por 693 descriptores moleculares. Las moléculas se clasifican en dos grupos: 226 moléculas tóxicas y 35 no tóxicas o seguras.

El conjunto de datos a utilizar como entrenamiento posee dos características que obligan a continuar realizando acciones como parte de la etapa de preprocesamiento. Estos datos poseen una alta dimensionalidad con respecto a la cantidad de instancias que posee, lo que propicia problemas de sobreajuste en la clasificación y un desbalance entre la cantidad de instancias que conforman los dos grupos en los que se clasifican estas moléculas teniendo un índice de desbalance  $IR = 6.45$  determinado mediante la razón entre las ocurrencias de la clase mayoritaria y las ocurrencias de la clase minoritaria. La realización de un entrenamiento con un conjunto de datos con estas características provoca que el criterio de los clasificadores sea parcializado a favor de las moléculas más representadas.

### **Selección de atributos, balanceo de datos y modelación**

Los softwares que calculan los descriptores moleculares generan entre cientos y miles de estos descriptores. En muchas ocasiones una gran parte de estos descriptores es información redundante y poco útil en el proceso de extracción del conocimiento. Por tal razón se hace necesario seleccionar aquellos atributos que son relevantes para el problema en cuestión y de esta forma se evita el riesgo de que los clasificadores se vean afectados por sobreajuste. La selección de atributos relevantes se realiza utilizando 3 enfoques: algoritmos filtros, algoritmos envolventes, algoritmos híbridos. En (Peng; Long y Ding 2005) se explican las características de los algoritmos que perteneces a estas tres clasificaciones. Para encaminar la solución del problema se utilizó el software mRMR (Ramentol y cols. 2011), que implementa una selección de atributos usando un criterio de dependencia estadística máxima basado en información mutua. El software permite determinar el subconjunto de atributos que posee una máxima relevancia con el atributo dependiente y a la vez poca redundancia entre ellos. La configuración de los parámetros para lograr estos resultados es especificando el tamaño

de subconjunto a obtener, el criterio de selección, el valor de corte de discretización (0.0, 0.5, 1.0) y el método de selección para la combinación de relevancia y redundancia. Los parámetros específicos para la corrida del software fueron: 50 atributos como tamaño del subconjunto, mRMR como criterio de selección, 0.0 como valor de corte de discretización (sin discretización) y MIQ como método de combinación de relevancia y redundancia.

Una vez realizada la selección de los atributos relevantes se realiza el balance de datos con el objetivo de modificar la distribución de los datos y garantizar que los resultados de los clasificadores no sean influenciados por el desbalance presente en los datos.

En conjuntos de entrenamientos desbalanceados la información de interés se encuentra en la clase minoritaria, muchos clasificadores consideran los datos poco representados como ruido o rarezas (Chawla y cols. 2002). En este trabajo se utilizó una técnica para balancear los datos que serán utilizados para realizar el entrenamiento, llamada *Synthetic Minority Oversampling Technique* (SMOTE) (Ramentol y cols. 2011).

SMOTE utiliza cada instancia de la clase minoritaria e introduce instancias sintéticas en el segmento que une a una/todas las instancias las k instancias más cercanas pertenecientes a la clase minoritaria. Dependiendo de la cantidad de nuevas instancias requeridas, los vecinos utilizados que pertenecen a los k vecinos más cercanos son seleccionados aleatoriamente.

Para la aplicación de este algoritmo se utilizó la implementación que está en la herramienta Weka (Breiman 1996).

Al finalizar estas actividades de preprocesamiento el conjunto de datos de entrenamiento queda reducido a un subconjunto de 50 atributos y presenta un IR=1.

Una vez obtenido el conjunto de datos de entrenamiento definitivo se procede a obtener modelos basados en multclasificadores. Para la obtención de los modelos, en una primera etapa, se evaluaron los ocho clasificadores mostrados en la Tabla 1, todos implementados en Weka.

<b>Grupo</b>	<b>Clasificadores</b>	<b>Abreviatura</b>
Arboles de decisión	ExtraTree	ET
Red bayesiana	NaiveBayes	NB
Basado en instancias (perezoso)	Kstar	K
Reglas	Jrip	J
Análisis discriminante	FDLA	AD
Regresión logística	Logistic	L
Red neuronal	MultilayerPerceptron	MLP
Máquina de soporte vectorial	SMO	SMO

Tabla 1: Clasificadores evaluados en la primera etapa de la obtención de los modelos

A continuación, se determinó el clasificador individual que mejor desempeño posee para ser utilizado como clasificador de control en la evaluación de los multclasificadores a utilizar en la experimentación. Los multclasificadores utilizados fueron Baggin, Boosting, Stacking y Vote.

Baggin (Freund y Schapire 1999) y Boosting (Wolper 1992) son algoritmos que garantizan diversidad en su aprendizaje debido a la estrategia utilizada para realizar el entrenamiento. El método de combinación de las salidas generadas es la utilización del voto mayoritario. En este trabajo se utilizó el algoritmo AdaboostM1 como multclasificador Boosting. Stacking (G. Casas) y Vote (Kuncheva 2007), a diferencia de los algoritmos mencionados anteriormente, no garantizan la diversidad en su aprendizaje. Para la emisión de un criterio sobre la clase de una instancia, estos algoritmos se basan en las salidas emitidas por los clasificadores utilizados en la construcción del modelo. En (Cabrera y Casas 2013) se especifican los mecanismos para garantizar la diversidad en la combinación de clasificadores a utilizar en la construcción de un multclasificador. Estos mecanismos son denominados “medidas de diversidad”.

En la realización de este trabajo se utilizó cinco medidas de diversidad: coeficiente de correlación ( $\rho$ ), estadístico Q, medida de diferencia (D), medida de doble fallo (DF) y razón entre D y DF denominado como (R). Para la realización de estos cálculos en (Cabrera y Casas 2013) plantea el procedimiento a realizar. Primeramente, hay que determinar las posibles combinaciones de clasificadores a utilizar. La determinación de las combinaciones de clasificadores más diversas se realizó mediante el cálculo de las medidas de diversidad para el cual se utilizó el software “Diversidad” desarrollado en (García; Luengo y Herrera 2008). Esta operación permitió determinar que un total de 32 combinaciones expresaban mayor diversidad ante el resto de las 216.

Las combinaciones diversas se organizaron de forma tal que permitiera realizar una experimentación con aquellas combinaciones que eran diversidad en las cinco medidas calculadas y otra experimentación con aquellas combinaciones que eran diversas en cuatro de las cinco medidas utilizadas. El algoritmo Stacking utiliza como método de combinación de las salidas de los clasificadores un clasificador, llamado metaclasificador, que entrena con un conjunto de instancias que se generan, donde cada componente de una instancia es la salida que emite cada clasificador individual utilizado en la construcción del modelo y el atributo dependiente de esta instancia es la clase real. En este caso se utilizó como metaclasificador el clasificador individual que mejor desempeño tuvo en la evaluación individual que se realizó.

El algoritmo Vote propone cinco métodos de combinación de las salidas de los clasificadores individuales, tales como: voto mayoritario, promedio de probabilidades, máximo de probabilidades, menor probabilidad y mediana de probabilidades. En este caso se utilizarán los métodos: voto mayoritario y promedio de probabilidades. Por tal razón la

experimentación deberá ser realizada sobre cada uno de los dos grupos de combinaciones de clasificadores construidas anteriormente utilizando para ellos los métodos de combinación anteriormente establecidos.

Finalmente se evaluaron los modelos obtenidos y se determinaron aquellos modelos que superaron en desempeño del mejor clasificador individual demostrando que se obtuvo una mejora en el desempeño respecto a los desempeños obtenidos con los clasificadores por separado.

Para la evaluación del desempeño de un clasificador se deben tener en cuenta la existencia de métricas que permiten evaluar cuan bien se comporta un modelo frente a una situación desconocida, es decir, permite evaluar la capacidad generalizadora de esos modelos. Existen varias métricas tales como desempeño. En este trabajo se utilizó el ACC, AUC y el estadístico Kappa, para evaluar el desempeño de los multclasificadores respecto al clasificador de control.

## **Resultados y discusión**

Luego de aplicar el procedimiento propuesto por Tropsha y colaboradores se eliminaron los compuestos que poseían elementos metálicos, organometálicos y poco representados.

Se utilizó ISIDA como herramienta para la determinación de los descriptores moleculares de las moléculas a partir de las cuales se procedió a realizar la experimentación. Se obtuvieron un total de 693 atributos por cada instancia del conjunto de datos.

A continuación, se utilizó el *ChemAxon's Standardizer* para realizar la estandarización de las moléculas y así evitar moléculas con un mismo grupo funcional, pero representadas con patrones estructurales diferentes y dar paso a la identificación de moléculas duplicadas. Para la identificación y posterior eliminación de duplicados se utilizó el módulo *ChemAxon's EdiSDF*. En este caso no se encontraron duplicados en el conjunto de datos.

Tropsha propone además que para la realización de estudios QSAR utilizando minería de datos el conjunto de datos debe ser particionado de forma que el 75% de las instancias pertenecientes a ambas clases sean utilizadas para realizar el entrenamiento de los clasificadores y el estante 25% de ellas sea utilizado como método de validación del aprendizaje inducido por los clasificadores. Luego de realizada esta operación se obtuvieron los conjuntos de datos a utilizar.

### **Selección de atributos, balanceo de datos y modelación**

Para llevar a cabo la selección de atributos se utilizó el software mRMR, que permitió la obtención de un subconjunto de 50 atributos relevantes y poco redundantes entre ellos.



A continuación, se procedió a realizar el balance de los datos. Mediante el algoritmo SMOTE, el cual permitió la creación de 143 instancias artificiales que garantizó un IR=1 en el conjunto de datos a utilizar para la obtención de los modelos.

Una vez balanceados los datos se realizó una evaluación con ocho clasificadores individuales, tal como se muestra en la Tabla 1. En la evaluación de cada clasificador individual se puede apreciar el clasificador *ET* con mejor desempeño (teniendo en cuenta las medidas ACC, AUC y KAPPA) y por lo tanto será utilizado para la evaluación de los modelos multclasificadores determinados en el proceso que sigue a continuación.

Se prosiguió con la evaluación de los multclasificadores Baggin y Boosting. Los resultados se pueden consultar en la Tabla 2. Luego de un análisis se determinó un modelo Boosting basado en una red neuronal (*MLP*) que supera al mejor clasificador individual evaluado.

Clasificador	Resultados modelos Baggin			Resultados modelos Boosting		
	ACC	AUC	Kappa	ACC	AUC	Kappa
ET	0.954	0.91	0.94	0.95	0.978	0.98
NB	0.924	0.96	0.95	0.95	0.98	0.98
K	0.954	0.96	0.96	0.95	0.976	0.97
J	0.924	0.97	0.96	0.95	0.988	0.98
revAD	0.924	0.94	0.94	0.90	0.955	0.96
L	0.893	0.93	0.93	0.90	0.951	0.96
<b>MLP</b>	0.939	0.98	0.98	<b>0.96</b>	<b>0.991</b>	<b>0.98</b>
SMO	0.924	0.95	0.94	0.90	0.967	0.96

Tabla 2: Resultados de los modelos Baggin y Boosting

Otros dos multclasificadores fueron evaluados, tal como se especificó anteriormente. Estos multclasificadores a diferencia de los anteriores no garantizan la diversidad en su aprendizaje por lo que es necesario que el usuario garantice la diversidad en la combinación de clasificadores que son utilizados en la obtención del modelo.

Se procedió a determinar las combinaciones posibles a obtener con los ocho clasificadores lo cual arrojó 248 combinaciones posibles a evaluar. Las medidas de diversidad fueron calculadas y solo se seleccionaron aquellas combinaciones de clasificadores que eran diversas en las cinco medidas de diversidad elegidas para este trabajo y otras combinaciones que fueron diversas en cuatro de las cinco medidas de diversidad calculadas. Esta información puede ser consultada en la Tabla 3 y la Tabla 4.

R	p	D	Q	DF	Subconjunto de Clasificadores
2.00	0.47	0.06	0.94	0.03	J—MLP
2.00	0.49	0.06	0.95	0.03	ET—SMO
2.00	0.49	0.06	0.95	0.03	AD—K

2.00	0.49	0.06	0.95	0.03	ET—AD
2.00	0.49	0.06	0.95	0.03	ET—NB
2.50	0.41	0.07	0.91	0.03	J—SMO
2.50	0.41	0.07	0.91	0.03	AD—J
2.50	0.41	0.07	0.91	0.03	J—NB
2.00	0.49	0.06	0.95	0.03	K—NB
2.00	0.49	0.06	0.95	0.03	K—SMO
1.61	0.59	0.05	0.95	0.03	ET--J—SMO
1.61	0.59	0.05	0.95	0.03	ET--J—NB
1.61	0.59	0.05	0.95	0.03	ET--AD—J
1.61	0.59	0.05	0.95	0.03	AD--J—K
1.61	0.59	0.05	0.95	0.03	J--K—SMO
1.61	0.59	0.05	0.95	0.03	J--K—NB
1.64	0.57	0.05	0.95	0.04	ET--AD--J—NB
1.64	0.57	0.05	0.95	0.04	AD--J--K—NB
1.64	0.57	0.05	0.95	0.04	ET--J--NB—SMO
1.64	0.57	0.05	0.95	0.04	J--K--NB—SMO

Tabla 3: Combinaciones más diversas en las cinco medidas de diversidad

R	p	D	Q	DF	Subconjunto de Clasificadores
1.83	0.53	0.06	0.94	0.04	J--NB—SMO
1.83	0.53	0.06	0.94	0.04	AD--J—NB
1.53	0.57	0.06	0.95	0.04	J--L—NB
1.58	0.59	0.05	0.95	0.04	AD--J—MLP
1.67	0.61	0.05	0.94	0.04	AD--J—SMO
1.58	0.59	0.05	0.95	0.04	J--MLP—SMO
1.58	0.59	0.05	0.95	0.04	J--MLP—NB
1.56	0.61	0.05	0.95	0.04	ET--AD--J—SMO
1.56	0.61	0.05	0.95	0.04	AD--J--K—SMO
1.48	0.61	0.05	0.96	0.04	AD--J--K--NB—SMO
1.48	0.61	0.05	0.96	0.04	ET--AD--J--NB--SMO

Tabla 4: Combinaciones más diversas en cuatro de las cinco medidas de diversidad

Una vez determinadas las combinaciones de clasificadores más diversas se procedió a realizar la experimentación utilizando el *Stacking* con las dos agrupaciones de combinaciones de clasificadores. Los resultados obtenidos en la experimentación fueron recogidos en las Tabla 5 y Tabla 6.

Subconjunto de Clasificadores	ACC	AUC	Kappa
J—MLP	0.92	0.909	0.717
ET—SMO	0.92	0.877	0.717
AD—K	0.90	0.854	0.647
ET—AD	0.93	0.918	0.764

ET—NB	0.95	0.88	0.797
J—SMO	0.92	0.805	0.692
AD—J	0.90	0.854	0.647
J—NB	0.92	0.909	0.717
K—NB	0.90	0.854	0.647
K—SMO	0.93	0.871	0.741
ET--J—SMO	0.90	0.851	0.614
ET--J—NB	0.89	0.845	0.605
ET--AD—J	0.92	0.863	0.692
AD--J—K	0.95	0.88	0.797
J--K—SMO	0.92	0.909	0.717
J--K—NB	0.92	0.863	0.692
ET--AD--J—NB	0.90	0.854	0.647
AD--J--K—NB	0.93	0.871	0.742
ET--J--NB—SMO	0.87	0.883	0.598
J--K--NB—SMO	0.95	0.927	0.815

Tabla 5: Resultados del desempeño de las combinaciones de clasificadores en las cinco medidas de diversidad con el multclasificador Stacking

Subconjunto de Clasificadores	ACC	AUC	Kappa
J--NB—SMO	0.863	0.921	0.913
AD--J—NB	0.909	0.901	0.915
J--L—NB	0.863	0.874	0.895
AD--J—MLP	0.893	0.84	0.887
AD--J—SMO	0.909	0.854	0.898
J--MLP—SMO	0.909	0.854	0.898
J--MLP—NB	0.909	0.854	0.898
ET--AD--J—SMO	0.893	0.845	0.89
AD--J--K—SMO	0.924	0.909	0.923
<b>AD--J--K--NB—SMO</b>	<b>0.969</b>	<b>0.936</b>	<b>0.956</b>
ET--AD--J--NB--SMO	0.863	0.874	0.895

Tabla 6: Resultados del desempeño de las combinaciones de clasificadores en cuatro de las cinco medidas de diversidad con el multclasificador Stacking

Como se puede apreciar en la Tabla 5 ninguna de las combinaciones posee un desempeño superior al mejor clasificador individual por tal razón no serán tenidas en cuenta para la selección del mejor multclasificador. En este caso el mejor resultado iguala al desempeño del mejor clasificador individual. En la Tabla 6 se encontró una combinación que supera en desempeño al mejor clasificador individual, por lo que estamos en presencia del segundo modelo a utilizar como solución al problema.

Luego de realizar un análisis se pudo determinar que los resultados previos demuestran que la forma de determinar diversidad reportada en la literatura no está correlacionada con una mejora en el desempeño del multclasificador utilizado. En este caso se identificó una combinación de clasificadores que según las medidas de diversidad no era muy diversa y sin embargo garantizó, con su utilización, una mejora en el desempeño del multclasificador utilizado permitiendo obtener un desempeño superior al desempeño del mejor clasificador individual utilizado para la evaluación de los multclasificadores.

Se procedió a realizar la experimentación ahora con el algoritmo Vote para la construcción del modelo. A continuación, se muestran los resultados de la experimentación con el multclasificador Vote para los dos grupos de clasificadores diversos registrados en las Tablas 7 y 8. Donde se utilizan los dos métodos de combinación de clasificadores propuestos anteriormente.

Método de combinación de Clasificadores	Promedio de probabilidades			Voto mayoritario		
	ACC	AUC	Kappa	ACC	AUC	Kappa
J—MLP	0.924	0.975	0.7647	0.924	0.816	0.6626
ET—SMO	0.939	0.980	0.7427	0.924	0.863	0.6927
AD—K	0.924	0.965	0.6927	0.924	0.863	0.6927
ET—AD	0.952	0.969	0.7975	0.924	0.863	0.6927
ET—NB	0.954	0.98	0.7975	0.924	0.98	0.7975
J—SMO	0.924	0.928	0.7179	0.907	0.807	0.614
AD—J	0.939	0.969	0.7161	0.909	0.807	0.614
J—NB	0.924	0.927	0.7179	0.909	0.807	0.614
K—NB	0.924	0.975	0.7179	0.924	0.863	0.6927
K—SMO	0.924	0.975	0.7179	0.924	0.863	0.6927
ET--J—SMO	0.954	0.978	0.7975	0.954	0.88	0.7975
ET--J—NB	0.954	0.978	0.7975	0.954	0.88	0.7975
ET--AD—J	0.954	0.967	0.7975	0.954	0.88	0.7975
AD--J—K	0.939	0.965	0.7161	0.954	0.88	0.7975
J--K—SMO	0.924	0.975	0.6927	0.954	0.88	0.7975
J--K—NB	0.924	0.975	0.6927	0.954	0.88	0.7975
ET--AD--J—NB	0.954	0.977	0.7975	0.939	0.871	0.7427
AD--J--K—NB	0.924	0.975	0.6927	0.939	0.871	0.7427
ET--J--NB—SMO	0.939	0.978	0.7647	0.939	0.871	0.7427
J--K--NB—SMO	0.939	0.975	0.7647	0.939	0.871	0.7427

Tabla 7: Resultados del desempeño de las combinaciones de clasificadores en las cinco medidas de diversidad con el multclasificador Vote por promedio de probabilidades y voto mayoritario

Método de combinación	Promedio de probabilidades			Voto mayoritario		
Subconjunto de Clasificadores	ACC	AUC	Kappa	ACC	AUC	Kappa
J--NB—SMO	0.939	0.928	0.7647	0.939	0.918	0.7647
AD--J—NB	0.939	0.971	0.7647	0.939	0.918	0.7647
J--L—NB	0.924	0.973	0.6927	0.924	0.963	0.6927
AD--J—MLP	0.939	0.975	0.7647	0.939	0.918	0.7647
AD--J—SMO	0.924	0.969	0.7179	0.924	0.909	0.7179
J--MLP—SMO	0.939	0.975	0.7647	0.939	0.918	0.7647
J--MLP—NB	0.939	0.973	0.7647	0.939	0.918	0.7647
ET--AD--J—SMO	0.954	0.977	0.9818	0.924	0.863	0.6927
AD--J--K—SMO	0.924	0.975	0.6927	0.939	0.863	0.6927
AD--J--K--NB—SMO	0.939	0.975	0.7647	0.939	0.918	0.7647
ET--AD--J--NB--SMO	0.939	0.977	0.7647	0.939	0.764	0.918

Tabla 8: Resultados del desempeño de las combinaciones de clasificadores en cuatro de las cinco medidas de diversidad con el multclasificador Vote por promedio de probabilidades y voto mayoritario

Como resultado del análisis de la información registrada en las Tablas 8 y 9 no se identificó ninguna combinación que superar al desempeño del mejor clasificador individual, por tal razón no serán tenidas en cuenta para la selección del mejor multclasificador

### Selección de los modelos para la solución del problema.

A partir de análisis de los resultados de los multclasificadores se identificaron dos multclasificadores que mostraron un desempeño superior al mejor desempeño logrado por los clasificadores individuales. Los resultados se compilan en la Tabla 9.

A partir de análisis realizado con los dos multclasificadores candidatos se pudo determinar que ambos multclasificadores no presentan diferencias significativas en su desempeño ya que poseen el mismo valor de la efectividad y una diferencia no significativa en los parámetros referentes al área bajo la curva ROC y el estadístico Kappa.

Por tal razón ambos multclasificadores constituyen las propuestas a utilizar para la clasificación de los LI (seguros o tóxicos) ante la posible inhibición del funcionamiento de la enzima AChE.

	ACC	AUC	Kappa
<b>Mejor clasificador individual</b>			
ET	0.954	0.928	0.7777
<b>Multclasificador Ada BoostM1</b>			
MLP	0.969	0.991	0.988
<b>Multclasificador Stacking</b>			
AD--J--K--NB—SMO	0.969	0.936	0.956

Tabla 9: Resultados de los mejores multclasificadores y el mejor clasificador individual

## Conclusiones

Con la realización de este trabajo se obtuvo un conjunto de datos referente a los lípidos iónicos y su efecto sobre la enzima AChE. Se aplicó, como parte del preprocesamiento a realizar sobre los datos, la limpieza manual propuesta por Tropsha a partir de la cual se eliminaron 28 moléculas. A continuación, se procedió a calcular los descriptores moleculares, se estandarizaron y se procedió a la identificación de duplicados en el conjunto de datos. Este último paso permitió determinar que no existían moléculas duplicadas en los datos. A continuación, se seleccionó el 75% de cada instancia para generar el conjunto de datos entrenamiento y el restante 25% fue utilizado para evaluar el aprendizaje. El software utilizado para el cálculo de los descriptores moleculares determinó 693 atributos por cada una de las instancias. A partir de esta situación se procedió a reducir la dimensionalidad del conjunto de datos, para ello se utilizó el software mRMR que permitió la obtención de un subconjunto de 50 atributos. A continuación, se aplicó el algoritmo SMOTE para modificar la distribución de los datos y de esta forma balancear la ocurrencia de ambas clases. Se procedió con la evaluación de ocho clasificadores que fueron utilizados como bases de los multclasificadores a utilizar en la obtención del modelo que prediga si un LI es tóxico o seguro ante la AChE. Luego de evaluados los clasificadores individuales, se procedió a evaluar los modelos obtenidos con los multclasificadores Baggin y Boosting y al compararlos con los clasificadores individuales, se determinó que el algoritmo AdaboostM1, perteneciente a la categoría de multclasificador Boosting, superó el desempeño del mejor clasificador individual utilizando como clasificador base una red neuronal MLP. Se procedió a evaluar los multclasificadores Stacking y Vote, a diferencia de Baggin y Boosting estos algoritmos no tienen en cuenta la diversidad de los clasificadores individuales a utilizar en la construcción del modelo, por tal razón el usuario debe determinar las combinaciones de clasificadores diversas. Para ello se calcularon

cinco medidas de diversidad reportadas en la literatura por cada combinación de clasificadores y luego de seleccionar las combinaciones más diversas se determinó otro modelo que superó al mejor clasificador individual, modelo conformado por el multclasificador Stacking y utilizando como clasificadores de base la combinación AD, J, K, NB y SMO. Por último, se compararon los dos modelos encontrados y se determinó que entre ellos no existen diferencias significativas por lo que se proponen ambos modelos para predecir si un líquido iónico desconocido si será capaz de inhibir la enzima AChE y por lo tanto provocar los efectos negativos mencionados en el organismo afectado. Lográndose darle cumplimiento al objetivo propuesto en el trabajo.

## Referencias

- ARNING, S. S. J. y cols. *Qualitative and quantitative structure activity relationships for the inhibitory effects of cationic head groups, functionalised side chains and anions of ionic liquids on acetylcholinesterase*, Green Chem, 2008. 10.
- ARNING, S. S. J. y A. BOSHEN. *Qualitative and quantitative structure activity relationship for the inhibitory effects of cationic head groups, functionalised side chains and anions of ionic liquids on acetylcholinesterase*, Green Chemistry, 2008. 10.
- BREIMAN, L. *Bagging predictors*, Machine Learning, 1996. 24.
- CABRERA, L. y G. CASAS. *Estudios de medidas de diversidad en sistemas multclasificadores*. Universidad de Las Villas. Martha Abreu, Laboratorio de Bioinformática, 2013.
- COULING, D. J. y cols. *Assessing the factors responsible for ionic liquid toxicity to aquatic organisms via quantitative structure–property relationship modeling*, Green Chem, 2006. 8.
- CHAWLA, N. y cols. *Smote: Synthetic minority over-sampling technique*, Artificial intelligence research, 2002. 16.
- CHEMAXON. "Chemaxon", 2012.
- CHEMAXON. "Standarizer," 2012.
- FRADE, C. A. A. R. F. y H. EXP., Toxicol, 2010. p.
- FREUND, Y. y R. SCHAPIRE. *A short introduction to boosting*, Journal of Japanese Society for Artificial Intelligence, 1999.

G. CASAS, J. L. M., L. DENODA, L. CABRERA, L. ARGUELLES, AND E. GONZALEZ, "High cardiovascular risk models for a city in the central region of cuba."

GARCÍA-LORENZO, E. T. A. y cols. *Cytotoxicity of selected imidazolium-derived ionic liquids in the human caco-2 cell line. Sub-structural toxicological interpretation through a qsar study*, Green Chem, 2008. 10.

GARCÍA, A. F. S.; L. LUENGO y F. HERRERA. *A study of statistical techniques and performance measures for genetics-based machine learning: Accuracy and interpretability*, Soft Comput, 2008. 13.

HOSSAIN, B. B. S. M. I. y cols. *Development of a novel mathematical model using a group contribution method for prediction of ionic liquid toxicities*, Chemosphere, 2011. 85.

ISIDA "Isida", 2008.

IZADIYAN, M. H. F. P. *In silico cytotoxicity estimation of ionic liquids based on their two and three-dimensional structural descriptors*, Monatsh Chem, 2011. 142.

J. KOOLMAN, A. K. H. R. *Color atlas of biochemistry*, Estados Unidos: Thieme, 2005.

J. P. J. S. TORRECILLA, J. L. A. F. R. —*a quantum-chemical-based guide to analyze/quantify the cytotoxicity of ionic liquids*. || Green Chem, 2010. p.

KUNCHEVA, L. I. *Combining pattern classifiers: Methods and algorithms (kuncheva, l.i.; 2004) [book review]*, Neural Networks IEEE Transactions on, 2007. 18.

LACRAMA, M. V. P. A. V. O. A.-M. *A spectral-sar model for the anionic-cationic interaction in ionic liquids: Application to vibrio fischeri ecotoxicity*, Int. J. Mol. Sci, 2007. 8.

LUIS, A. G. P. y A. IRABIEN. *Quantitative structure – activity relationships (qsars) to estimate ionic liquids ecotoxicity ec50 (vibrio fischeri)*, Molecular Liquids, 2010. 152.

LUIS, I. O. P.; R. ALDACO y A. IRABIEN. *A novel group contribution method in the development of a qsar for predicting the toxicity (vibrio fischeri ec50) of ionic liquids*, Ecotoxicol. Environ. Saf, 2007. 67.

MATHEWS., H., AND AHERN. *Biochemistry*, 2005.

METZLER, D. E. *Biochemistry. The chemical reactions of living cells*, ELSEVIER. Academic Press, 2003.

MOHAMMAD, P. I. y H. FATEMI. *Cytotoxicity estimation of ionic liquids based on their effective structural features*, Chemosphere, 2011. 84.



PENG, H.; F. LONG y C. DING. *Feature selection based on mutual information: Criteria of max-dependency, max-relevance and minredundancy*, IEEE Transactions on pattern recognition and machine intelligence, 2005.

PUTZ, A.-M. L. M. A. V. O. M. V. *Res. Lett. Ecology*, 2007. p.

RAMENTOL, E. y cols. *Smote-rs b\*: A hybrid preprocessing approach based on oversampling and undersampling for high*

*imbalanced data-sets using smote and rough sets theory*, Knowl Inf Syst, 2011.

SIERRA, A. G. J.; E. MARTÍ y R. CRUAÑAS. *Química verde: Toxicidad y biodegradabilidad de líquidos iónicos*. Universitat de Barcelona, Departament de Productes Naturals, Biologia Vegetal i Edafologia, Unitat d'Edafologia, 2006.

STOCK, F. y cols. *Effects of ionic liquids on the acetylcholinesterase -a structure-activity relationship consideration*, Green Chem, 2004. 6.

STÖRMANN, J. A. R. y cols. *Uft merck ionic liquids biological effects database*, 2014. [Disponible en: <http://www.ileco.uft.uni-bremen.de>.

TORRECILLA, J. S. y cols. *A quantum-chemical-based guide to analyze/quantify the cytotoxicity of ionic liquids*, Green Chemistry, 2010. 12.

TROPSHA, D. F. A. y E. MURATOV. *Trust, but verify: On importance of chemical structure curation in cheminformatics and qsar modeling research*, J. Chem. Inf. Model, 2010.

V. SOLOV'EP, A. A. V., . "Editor of the structure data file," 2010.

WOLPER, D. H. *Stacked generalization*, Neural Networks, 1992. vol. 5.

ZHAO, Y. L. D. y Z. ZHANG. *Toxicity of ionic liquids*, Clean Prod. Proces., 2007. 35.