



LINEAS DE PRODUCTOS SOFTWARE

GRUPO GLASS

20 DE ENERO DE 2020
ALEJANDRO FRANCISCO GARCIA UCLES
JUAN SOLER MÁRQUEZ
JOSE MANUAL MARTINEZ SALAS
DIEGO CANGAS MOLDES



UNIVERSIDAD DE ALMERÍA

Proyecto Glass

LINEA DE PRODUCTOS SOFTWARE



Escuela Superior de Ingeniería
UNIVERSIDAD DE ALMERÍA

Almería, 2 de diciembre de 2019

Sesión: Análisis de datos con Weka

Miembros:

Juan Soler Márquez

Diego Cangas Moldes

Alejandro Francisco García Uclés

José Manuel Martínez Salas

Índice

1. Introducción	3
2. Modelos caja blanca	3
2.1. J48	3
2.1.1. J48 Con todos los atributos	3
2.1.2. J48 Con selección de atributos	4
2.2. LMT	6
2.2.1. LMT con todos los atributos	6
2.2.2. LMT con selección de atributos	6
2.3. Conclusiones	7
3. Modelo caja negra	8
3.1. MLP	8
3.1.1. MLP con todos los atributos	8
3.1.2. MLP con selección de atributos	8
3.2. Conclusiones	9

1. Introducción

Para este análisis hemos utilizado el dataset Glass.arff. Este dataset se trata de un conjunto de datos sobre cristales. Contiene información sobre la composición química de cada uno de los cristales, el índice de refracción, la técnica con la que se ha conformado y la clase a la que pertenece.

2. Modelos caja blanca

2.1. J48

2.1.1. J48 Con todos los atributos

Primero hemos elegido el algoritmo J48 el cual arroja un 66,82 % de instancias clasificadas correctamente.

Para ello cargamos nuestro conjunto de datos *glass.arff* y luego pinchamos en *Classify*, una vez dentro pinchamos en *Choose* y dentro de *tree*, elegimos *J48* pulsamos *Start* y esperamos el resultado. El cual podemos verlo en la *Figura 2*.

<pre> === Run information === Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2 Relation: Glass Instances: 214 Attributes: 10 RI Na Mg Al Si K Ca Ba Fe Type Test mode: 10-fold cross-validation === Classifier model (full training set) === J48 pruned tree </pre>		<pre> Number of Leaves : 30 Size of the tree : 59 Time taken to build model: 0.06 seconds === Stratified cross-validation === === Summary === Correctly Classified Instances 143 66.8224 % Incorrectly Classified Instances 71 33.1776 % Kappa statistic 0.55 Mean absolute error 0.1826 Root mean squared error 0.2857 Relative absolute error 48.4587 % Root relative squared error 89.2727 % Total Number of Instances 214 === Detailed Accuracy By Class === TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class 0.714 0.174 0.667 0.714 0.690 0.532 0.806 0.667 build wind float 0.618 0.181 0.653 0.618 0.635 0.443 0.768 0.606 build wind non-float 0.353 0.046 0.400 0.353 0.375 0.325 0.766 0.251 vehic wind float ? 0.000 ? ? ? ? ? ? vehic wind non-float 0.769 0.010 0.833 0.769 0.800 0.788 0.872 0.575 containers 0.778 0.029 0.538 0.778 0.636 0.629 0.930 0.527 tableware 0.793 0.022 0.852 0.793 0.821 0.795 0.869 0.738 headlamps Weighted Avg. 0.668 0.130 0.670 0.668 0.668 0.539 0.807 0.611 === Confusion Matrix === a b c d e f g <-- classified as 50 15 3 0 0 1 1 a = build wind float 16 47 6 0 2 3 2 b = build wind non-float 5 5 6 0 0 1 0 c = vehic wind float 0 0 0 0 0 0 0 d = vehic wind non-float 0 2 0 0 10 0 1 e = containers 1 1 0 0 0 7 0 f = tableware 3 2 0 0 0 1 23 g = headlamps </pre>	
--	--	--	--

Figura 1: Atributos

Figura 2: Resultado

2.1.2. J48 Con selección de atributos

Después de ver los resultados que nos devuelve el J48 con todos los atributos, hemos aplicado una selección de atributos para ver cuáles eran más relevantes y cuáles no. Tras realizar la selección de atributos, nos quedamos con los atributos siguientes:

- Refractive Index
- Magnesium
- Potassium
- Aluminium
- Barium

Con estos atributos, el J48 arroja un 71,028 % de instancias clasificadas correctamente.

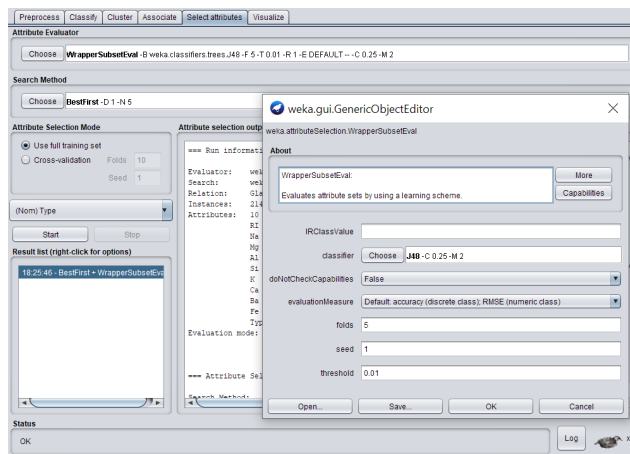


Figura 3: Selección de atributos

Para realizar esta operación, entramos en la pestaña *Select Attributes* de *Weka* y en *Attribute Evaluator* seleccionamos *WrapperSubSetEval* y dentro de este en *Classifier* seleccionamos el atributo con el que hemos hecho la clasificación anteriormente para ver qué atributos no son significativos y podemos eliminarlos de nuestro conjunto de datos. Así luego, ejecutaremos de nuevo el algoritmo de clasificación para compararlo con el resultado anterior con todos los atributos.

```

=== Run information ===

Scheme:      weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:    Glass-weka.filters.unsupervised.attribute.Remove-R2,5,7,9
Instances:   214
Attributes:  6
              RI
              Mg
              Al
              K
              Ba
              Type
Test mode:   10-fold cross-validation

=== Classifier model (full training set) ===

J48 pruned tree

Ba <= 0.27
|
|_ Mg <= 2.41
|   |_ K <= 0.03
|       |_ RI <= 1.52315: tableware (10.0/1.0)
|       |_ RI > 1.52315: build wind non-float (2.0)
|       |_ K > 0.03
|           |_ Al <= 1.38: build wind non-float (6.0/1.0)
|           |_ Al > 1.38
|               |_ Mg <= 1.88: containers (14.0/2.0)
|               |_ Mg > 1.88: build wind non-float (3.0/1.0)
|   |_ Mg > 2.41
|       |_ Al <= 1.41
|           |_ RI <= 1.51707
|           |   |_ RI <= 1.51596: build wind float (3.0)
|           |   |_ RI > 1.51596
|           |       |_ Al <= 1.27
|           |           |_ Mg <= 3.46: vehic wind float (3.0/1.0)
|           |           |_ Mg > 3.46: build wind non-float (3.0)
|           |           |_ Al > 1.27: vehic wind float (5.0)
|           |   |_ RI > 1.51707
|           |       |_ K <= 0.23
|           |           |_ Mg <= 3.34: build wind non-float (2.0)
|           |           |_ Mg > 3.34
|           |               |_ RI <= 1.52127
|           |                   |_ Al <= 0.91: vehic wind float (5.0/1.0)
|           |                   |_ Al > 0.91: build wind float (5.0/1.0)
|           |                   |_ RI > 1.52127: build wind float (15.0/1.0)
|           |       |_ K > 0.23
|           |           |_ Mg <= 3.75
|           |               |_ Al <= 1.16
|           |                   |_ RI <= 1.51806: build wind float (7.0/1.0)
|           |                   |_ RI > 1.51806: build wind non-float (5.0)
|           |                   |_ Al > 1.16: build wind float (37.0/2.0)
|           |                   |_ Mg > 3.75: build wind non-float (10.0)
|           |       |_ Al > 1.41: build wind non-float (50.0/10.0)
|   |_ Ba > 0.27: headlamps (29.0/3.0)

```

Figura 4: Atributos J48sin

```

Number of Leaves :    19
Size of the tree :    37

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances      152           71.028 %
Incorrectly Classified Instances    62           28.972 %
Kappa statistic                    0.6019
Mean absolute error                 0.1019
Root mean squared error             0.2754
Relative absolute error             48.0989 %
Root relative squared error        84.8641 %
Total Number of Instances         214

=== Detailed Accuracy By Class ===

```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
0,829	0,125	0,763	0,829	0,795	0,690	0,860	0,699	build wind float	
0,697	0,181	0,679	0,697	0,688	0,513	0,754	0,600	build wind non-float	
0,235	0,036	0,364	0,235	0,286	0,245	0,766	0,232	vehic wind float	
?	0,000	?	?	?	?	?	?	vehic wind non-float	
0,692	0,015	0,750	0,692	0,720	0,703	0,021	0,503	containers	
0,556	0,024	0,500	0,556	0,526	0,505	0,073	0,448	tableware	
0,793	0,022	0,852	0,793	0,821	0,795	0,063	0,634	headlamps	
Weighted Avg.	0,710	0,113	0,702	0,710	0,704	0,599	0,613	0,596	

```

=== Confusion Matrix ===
a b c d e f g <-- classified as
58 9 1 0 0 1 1 | a = build wind float
9 53 6 0 3 3 2 | b = build wind non-float
6 6 4 0 0 1 0 | c = vehic wind float
0 0 0 0 0 0 0 | d = vehic wind non-float
0 3 0 0 9 0 1 | e = containers
0 4 0 0 5 0 0 | f = tableware
3 3 0 0 0 0 23 | g = headlamps

```

Figura 5: Resultado J48sin

2.2. LMT

2.2.1. LMT con todos los atributos

El algoritmo LMT cuando lo ejecutamos usando todos los atributos arroja un resultado de 68.6916 %.

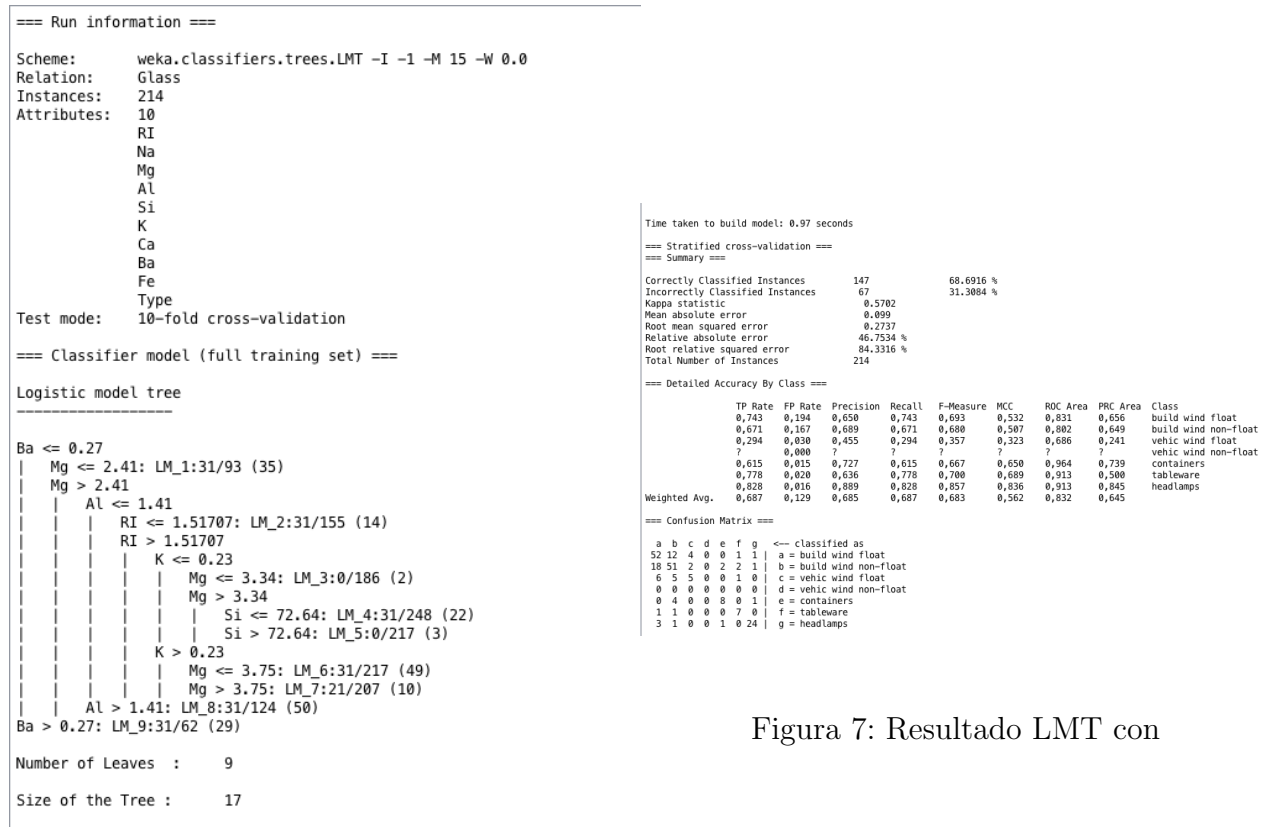


Figura 7: Resultado LMT con

Figura 6: Atributos LMT con

2.2.2. LMT con selección de atributos

Después de ver los resultados que nos devuelve el LMT con todos los atributos, hemos aplicado una selección de atributos para ver cuáles eran más relevantes y cuáles no. Tras realizar la selección de atributos, nos quedamos con los atributos siguientes:

- Refractive Index
- Magnesium
- Potassium
- Aluminium
- Barium
- Calcium

Con estos atributos, el LMT arroja un 66.8224 % de instancias clasificadas correctamente.

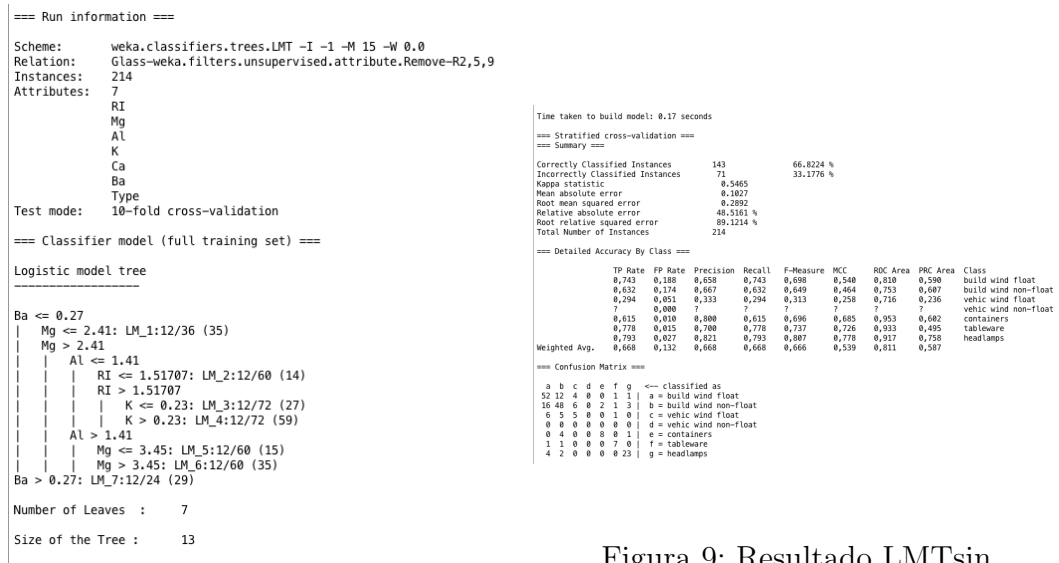


Figura 9: Resultado LMTsin

Figura 8: Atributos LMTsin

2.3. Conclusiones

Una vez analizados los resultados con los algoritmos J48 y LMT, hemos decidido usar el algoritmo J48 para implementar el algoritmo de caja blanca en nuestra aplicación. Las razones por la cual hemos elegido ésta es porque el porcentaje de instancias clasificadas correctamente es mejor y el árbol de decisión resultante es bastante más simple que el árbol que genera el algoritmo LMT.

3. Modelo caja negra

3.1. MLP

3.1.1. MLP con todos los atributos

El algoritmo MLP, de caja negra, cuando lo ejecutamos usando todos los atributos arroja un resultado de clasificación correcta de un 67.757%.

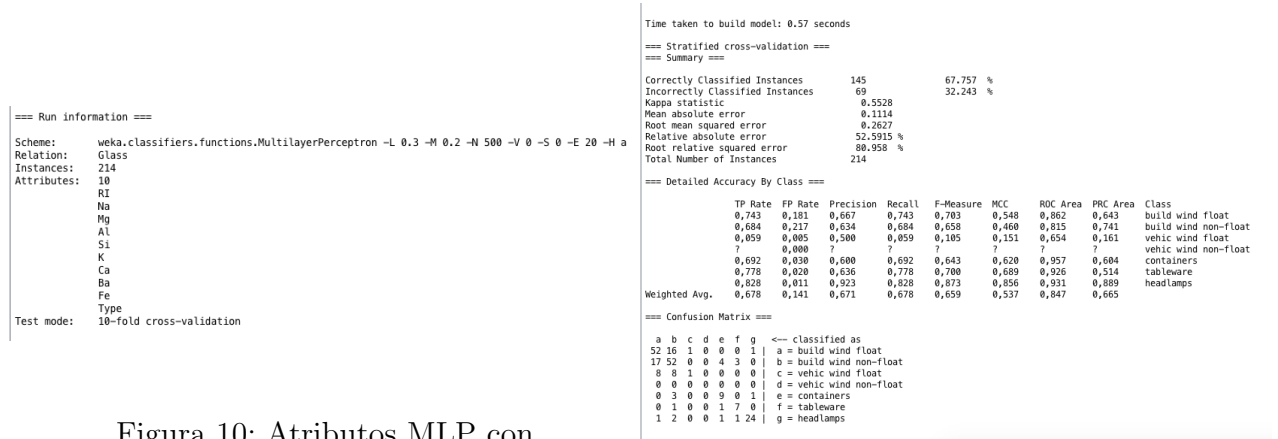


Figura 10: Atributos MLP con

Figura 11: Resultado MLP con

3.1.2. MLP con selección de atributos

Como podemos ver el resultado con la selección de atributos empeora, por lo tanto para nuestro modelo de caja negra, utilizaremos todos los atributos del conjunto de datos.

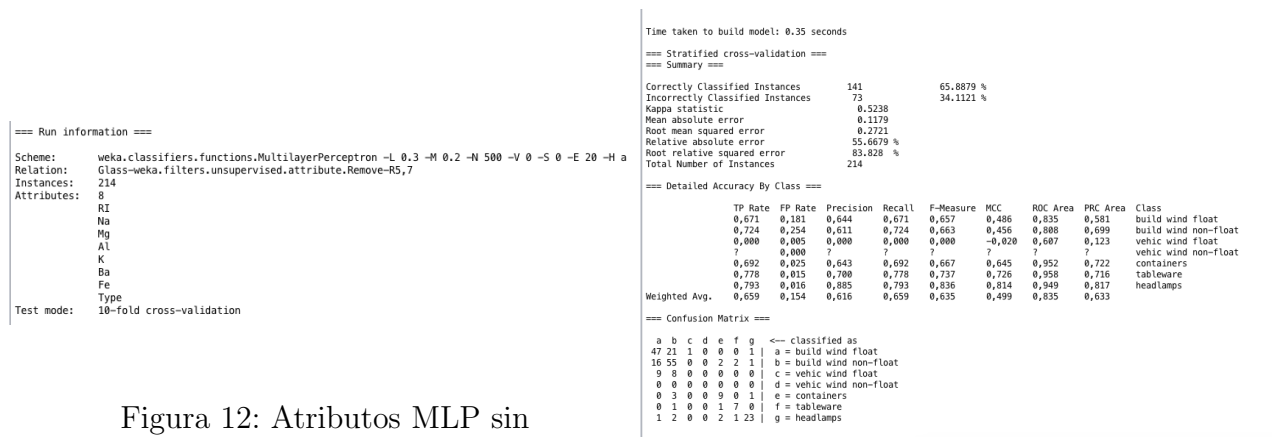


Figura 12: Atributos MLP sin

Figura 13: Resultado MLP sin

3.2. Conclusiones

Para nuestro modelo de caja negra, hemos decidido usar el algoritmo MLP con todos los atributos, ya que al hacer uso de la selección de atributos el porcentaje de instancias clasificadas correctamente disminuye. Por ello, en la interfaz de nuestra aplicación, cuando tengamos que introducir datos para algoritmo caja negra tendremos disponibles todos los atributos, en cambio, cuando introduzcamos valores para el algoritmo de caja blanca, solo aparecerán activos los atributos seleccionados en el paso 2.1.2

●●●○○ ABC

09:22 PM



Version 1.0

Email

Password

Registe

Sign In

●●●○○ ABC

09:23 PM



Logout

Cases



Case: **El caso del gallinero**

SampleNr: **15** Created: **17.1.2020**

ABC

09:31 PM



< Cassettes

Members

Save

joseJua@hotmail.com



FranccoFr@hotmail.com



Elyoni@hotmail.com



Administrador

●●●○○ ABC

09:33 PM



< Home

Cases

Add case

Case: **El caso del gallinero**

SampleNr: **15** Created: **17.1.2020**

ABC

12:26 AM



< Cases

Samples



Team Update

SampleCrystal1

CrystalType: Tableware

SampleCrystal2

CrystalType: Vehic wind non-floc

ABC

09:33 PM



< Cases

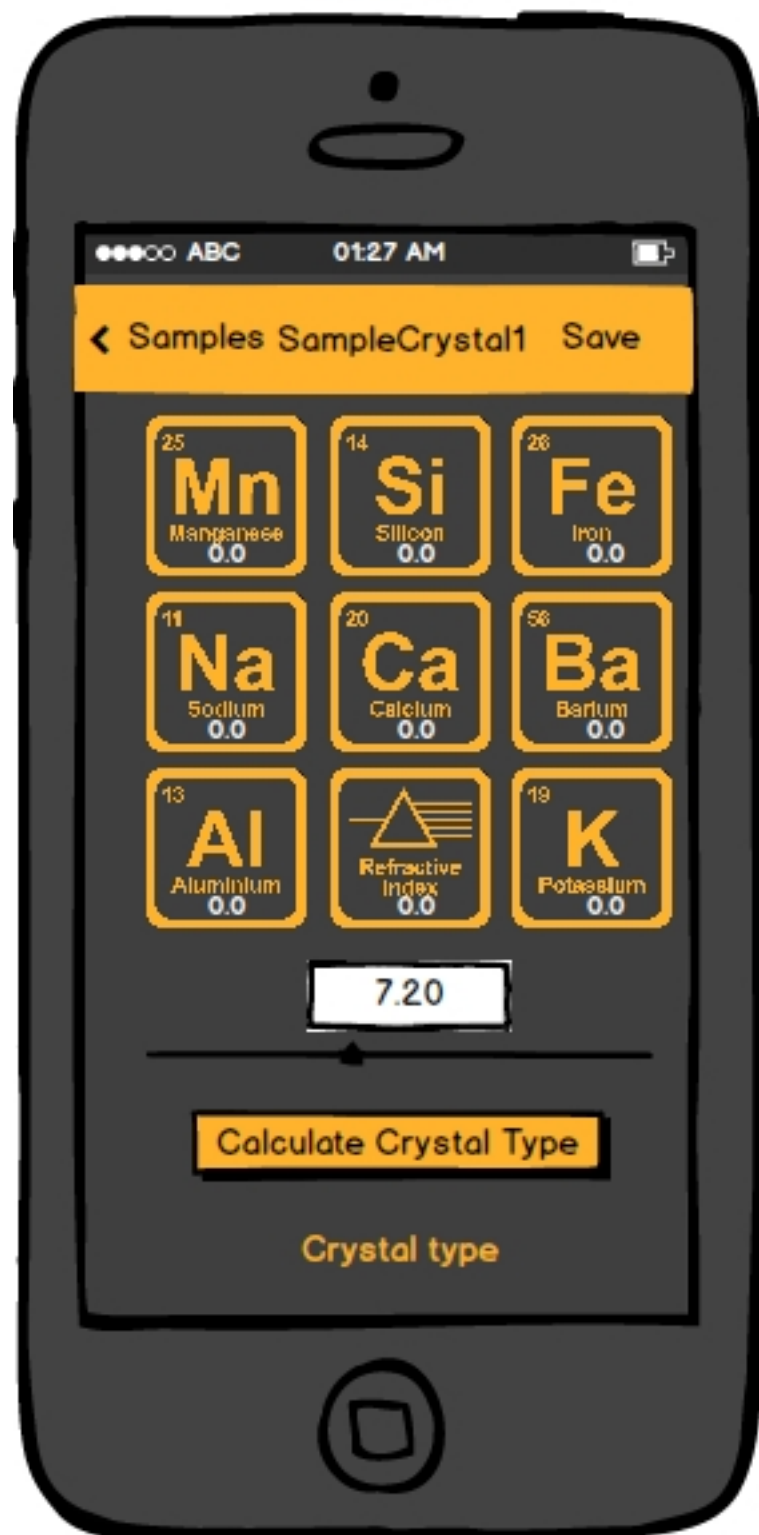
New case

Save

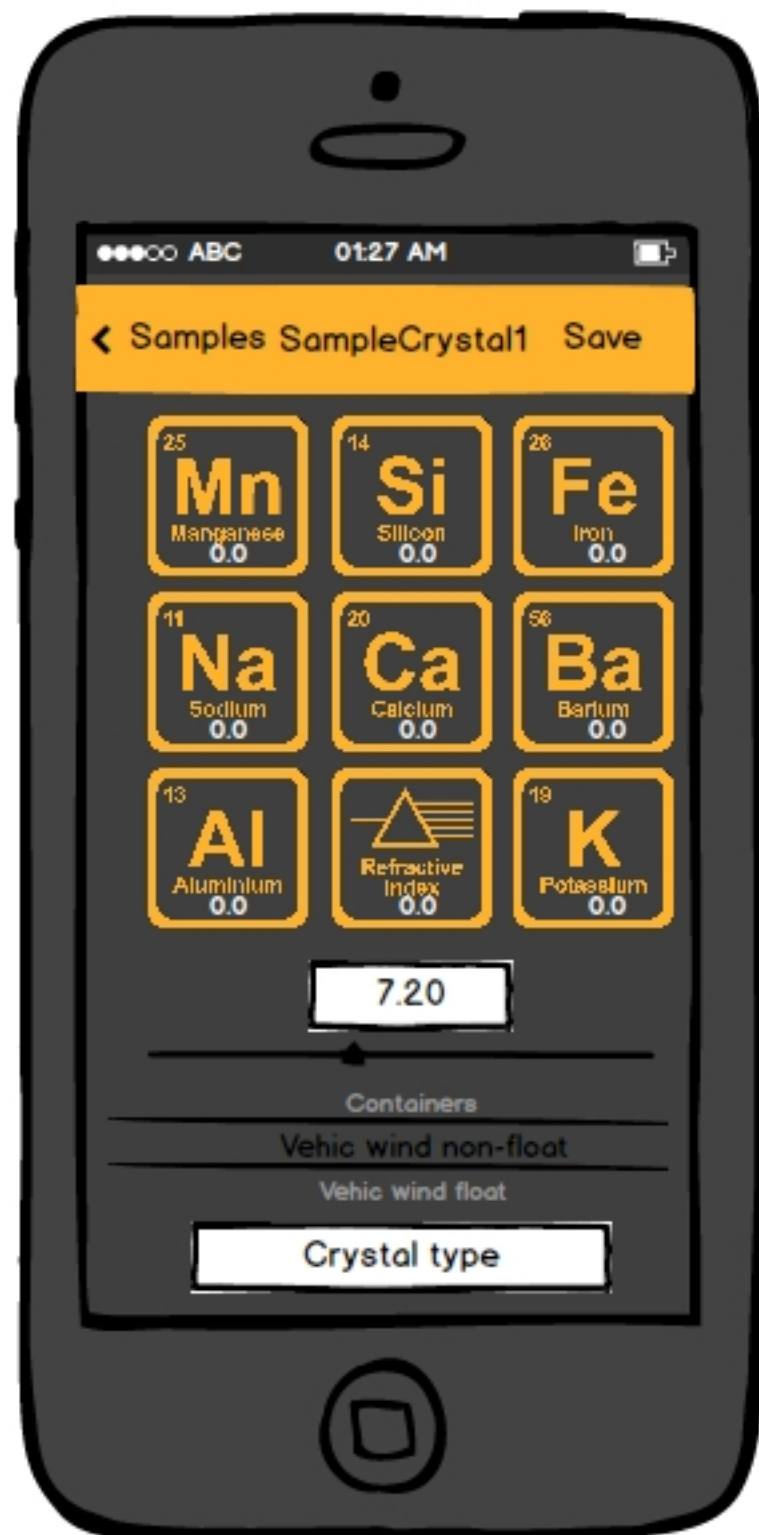
Case Name

Add members

Version 1.0



Version 2.0



●●●○○ ABC

09:44 PM



< Cases

Add member

Save

Email

Admin



●●●● ABC

09:31 PM



Logout

Admin

User Mode

Add Case

Add User

ABC

09:49 PM



Cancel

SignUp



Email

Password

ImageUp

SignUp

ABC

12:41 AM



< Cases Samples



Team Update

Sample Add

Cancel

Save

ABC

12:41 AM



< Cases

Add member

Save

Email

Admin



Success

Member was successfully created.

Done

ABC

12:41 AM



Cancel

SignUp



Email

**User was successfully
created.**

Done

Version 1.0



Version 2.0



Aspectos destacados de implementación

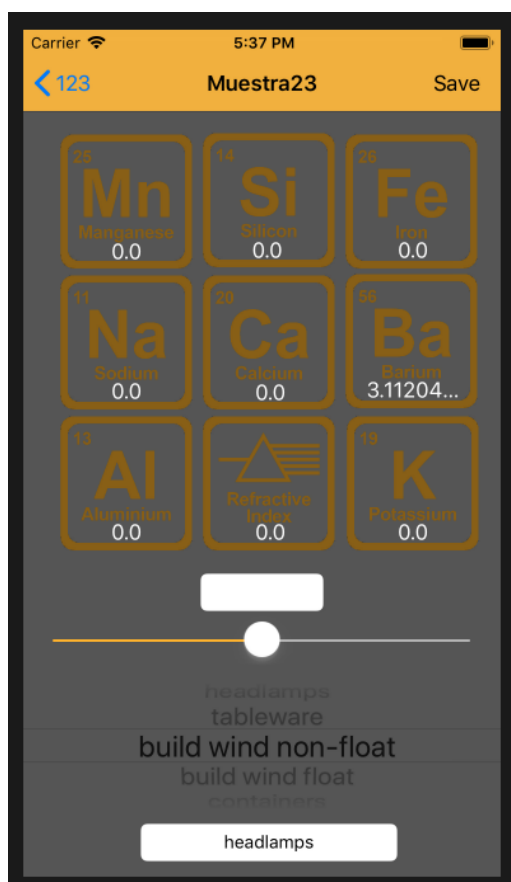
En este apartado se comentarán algunos aspectos destacados sobre la implementación de nuestra aplicación.

Esta aplicación tiene como objetivo facilitar la información sobre muestras en investigaciones sobre tipos de cristales a través de su composición química.

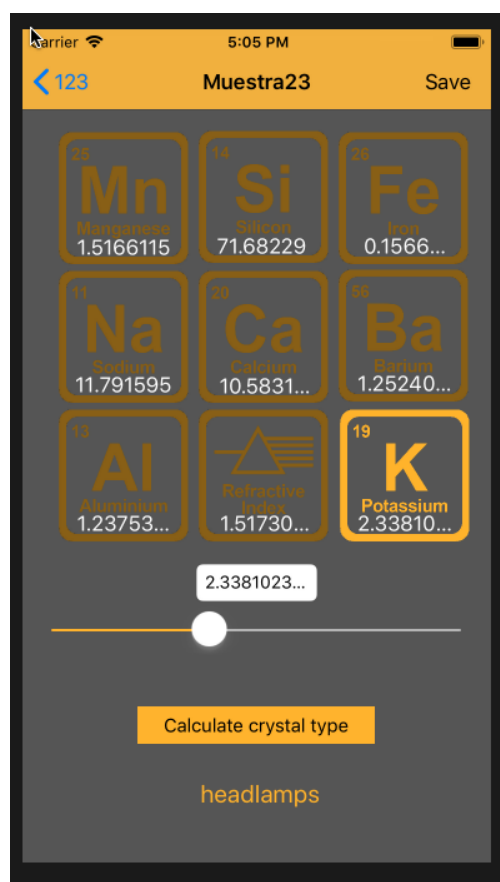
Esta aplicación está dividida en dos versiones 1.0 y 2.0. Las diferencias entre la versión 1.0 y 2.0 son que en la versión 1.0, el investigador elije el tipo de cristal al que pertenece la muestra mientras que en la versión 2.0 el tipo de cristal que es se calcula mediante su composición química.

Esto se lleva a cabo en la interfaz de detalles de muestra donde introduciendo los componentes esta misma te calcula el tipo de cristal que es.

Versión 1.0



Versión 2.0

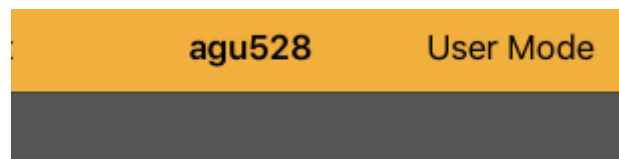


En esta ventana de la interfaz, cada elemento químico y el índice de refracción son un botón que, al pulsarlo, cambia el límite del slider(campo de desplazamiento) justo debajo del textfield (campo de texto) con el valor de ejemplo “7.20”. Esto se debe a que cada elemento químico para nuestro caso tendrá uno mínimo y un máximo distinto. El campo de texto que hemos pasado por encima antes, muestra el valor del elemento químico señalado ahora mismo, pudiéndose cambiar el valor de este campo de texto y afectando al mismo químico.

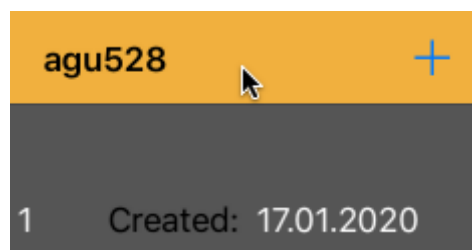
La barra deslizante tiene como objetivo facilitar la introducción de valores. Como dijimos antes, en la versión 1 elegiremos el tipo de cristal con nuestro picker, pero en la versión dos pulsando el botón “Calculate Crystal Type” nos devuelve el tipo de cristal justo en el campo de texto no editable de abajo.

Otro aspecto destacado en nuestra implementación es la posibilidad de que el administrador acceda al modo usuario. Desde este modo, el administrador puede llevar a cabo la adición de nuevos casos mientras que todavía puede ver lo mismo que ve un usuario.

Entrar a modo usuario



Añadir nuevo caso desde modo usuario



En nuestra aplicación, tenemos tres tipos de usuarios:

1. Usuario admin
2. Usuario miembro
3. Usuario no miembro

Estos tipos de usuario han sido listados jerárquicamente, de más poder en la aplicación a menos. La razón de que se añada en aspectos destacados reside en la posibilidad de no ser miembro.

¿Qué es un miembro? Un miembro es un usuario que ha sido validado por un usuario administrador por medio del “Add User”.

¿Por qué esta explicación? Todo el mundo puede registrarse en nuestra aplicación, pero solo aquellos que estén validados como miembros por un usuario administrador podrán ser parte de los casos. Y ya que para ver un caso, un usuario administrador debe haberte incluido en la lista de miembros de ese caso, nunca podrías ver caso alguno ya que ni siquiera podrías ser incluido en estos.

Nuestra aplicación cuenta también con un sistema que mantiene al usuario logueado aunque abandone la aplicación. Esto viene de parte de Firebase, un sistema de bases de datos en línea que hemos decidido usar para nuestra aplicación.

Y puesto que Firebase será nuestra base de datos, lo conveniente es dar una buena explicación.

Lo primero que necesita un proyecto Swift para que se conecte con Firebase es agregar los SDK de Firebase a nuestra App y eso se hace usando la herramienta CocoaPods, es parecido a npm. En nuestro archivo PodFile, agregamos los paquetes que queremos agregar a nuestro proyecto. Por ejemplo, si queremos usar el servicio de autenticación de Firebase debemos agregar la siguiente línea:

```
pod 'Firebase/Auth'
```

De este modo ya podremos hacer `import FirebaseAuth` y poder hacer uso de los métodos `Auth.auth().signIn` o `Auth.auth().createUser`. y poder loguearnos, crear usuarios, etc.

Es muy importante inicializar nuestra conexión con Firebase antes de hacer uso de cualquiera de sus servicios, para ello debemos agregar `FirebaseApp.configure()` en nuestro `AppDelegate` en `func application`, para que lo inicialice al arrancar nuestra aplicación.

Otro paso fundamental es agregar el archivo `GoogleService-Info.plist` el cual lo podemos descargar directamente desde el apartado de configuración de nuestro proyecto en Firebase. Este archivo contiene identificadores de proyecto de Firebase para enlazar nuestra aplicación a nuestro proyecto en Firebase.

Para hacer uso del servicio de base de datos de Firebase, tenemos que agregar el pod 'Firebase/Database', gracias a este import podremos llamar a `Database.database().reference(withPath: "nombre_de_la_coleccion")` y con ello ya podremos conseguir los hijos de la colección que queremos o bien crear hijos nuevos, actualizarlos o eliminarlos.

Con Firebase podemos modelar objetos para que sea mas fácil hacer CRUD en la base de datos, olvidándonos de crear sentencias personalizadas para actualizar algún atributo de nuestro objeto.

Para poder hacer esto, creamos un modelo e importamos `Firebase` y `FirebaseDatabase`, ahora ya podemos crear un constructor que pase una referencia de un objeto de ese tipo que este alojado en Firebase y lo convertirá automáticamente a un objeto compatible para que podamos usarlo y llamar a todos los atributos o métodos que hayamos definido en él. Obviamente, también tendrá su constructor clásico para poder inicializar un objeto pasándole los parámetros que hayamos definido en él (nombre, fecha, addedByUser, etc).

```
init(name: String, addedByUser: String, completed: Bool, key: String = "", fecha: String) {
    self.ref = nil
    self.key = key
    self.name = name
    self.addedByUser = addedByUser
    self.completed = completed
    self.fecha = fecha
}

init?(snapshot: DataSnapshot) {
    guard
        let value = snapshot.value as? [String: AnyObject],
        let name = value["name"] as? String,
        let addedByUser = value["addedByUser"] as? String,
        let fecha = value["fecha"] as? String,
        let completed = value["completed"] as? Bool else {
        return nil
    }

    self.ref = snapshot.ref
    self.key = snapshot.key
    self.name = name
    self.addedByUser = addedByUser
    self.completed = completed
    self.fecha = fecha
}
```

Para hacer la función inversa, convertir nuestro objeto en un objeto tipo JSON que es lo que necesita Firebase para agregar la información en la base de datos creamos este método en nuestro modelo.

```
func toAnyObject() -> Any {  
    return [  
        "name": name,  
        "addedByUser": addedByUser,  
        "completed": completed,  
        "fecha": fecha  
    ]  
}
```

En la siguiente imagen, podemos ver como creamos un objeto de tipo Case llamado caseItem. Luego creamos un hijo y guardamos su referencia en caseItemRef y finalmente agregamos el objeto Case a ese hijo con setValue y le pasamos el toAnyObject para que tenga la forma JSON deseada por Firebase.

```
let caseItem = Case(name: nameText.text!,  
    addedByUser: AppDelegate.user!.email,  
    completed: false, fecha: formatter.string(from:  
        date))  
  
let caseItemRef = Database.database().reference(withPath: "case-  
items").child(nameText.text!.lowercased())  
  
caseItemRef.setValue(caseItem.toAnyObject())
```

Si queremos ver si un hijo existe, podemos hacer uso del método observe, el cual recibe los valores que tenga ese hijo y podemos comprobar si existe, recorrerlo, etc.

En la siguiente imagen, comprobamos si existe y si existe ya podemos obtener su valor con `snapshot.value` o incluso si tiene hijo a su vez, con `snapshot.children`.

```
let ref = Database.database().reference(withPath:
    "members").child(items[0]).child("admin")
UserDefaults.standard.set(true, forKey: "loggedin")

ref.observe(.value, with: { snapshot in

    if !snapshot.exists() {
```

En la siguiente imagen vemos como recorremos los hijos de un hijo, en este caso, dentro de la colección “members” hay un hijo llamado “cases”. Como podemos ver, podemos pedirle a Firebase que nos devuelva a los hijos ordenados por algún atributo y puede ser ASC o DESC. Una vez que empezamos a recorrer a los hijos y obtenemos objetos de tipo `DataSnapshot`, recordamos el constructor que creamos en nuestro modelo el cual recibía como parámetro un `DataSnapshot` y que el cual pudimos establecerlo gracias al import `FirebaseDatabase` que pusimos en nuestra clase.

Al llamar a este constructor, automáticamente convierte el `Snapshot` en un objeto de tipo `Case` perfectamente valido para poder trabajar con él.

```
let ref = Database.database().reference(withPath: "members")

ref.child(AppDelegate.username!).child("cases").queryOrdered(
    byChild: "completed").observe(.value, with: { snapshot in
    var newItems: [Case] = []
    for child in snapshot.children {
        if let snapshot = child as? DataSnapshot,
            let caseItem = Case(snapshot: snapshot) {
            print(caseItem.name)
            newItems.append(caseItem)
        }
    }

    self.items = newItems
    self.tableView.reloadData()
})
```

Una de las funcionalidades a destacar de Firebase es el uso de Observables, gracias a ello, cualquier cambio que se realice en nuestra base de datos, se actualizará automáticamente en nuestra aplicación y todo ello “gratis” sin tener que llamar a listeners, valueChanged u otros métodos que estén atentos para ver si ha habido algún cambio en nuestra base de datos.