

SelAtt

Sistema de información enfocado en:
Machine learning aplicado a datos biomédicos

ANDRÉS CARRILLO BEJARANO

TUTORES:
FRANCISCO GOMÉZ VELA
MIGUEL GARCIA TORRES

ÍNDICE

- 1.Objetivos
- 2.Solución propuesta
- 3.Arquitectura
- 4.Tecnologías
- 5.Demo Live
- 6.Experimentación
- 7.Conclusiones

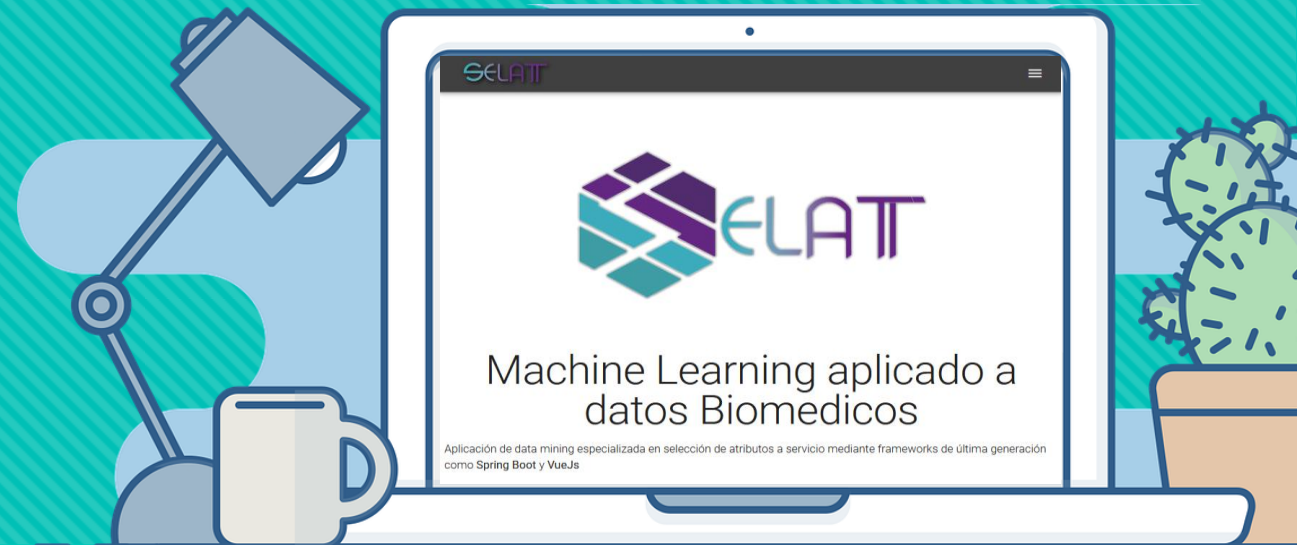
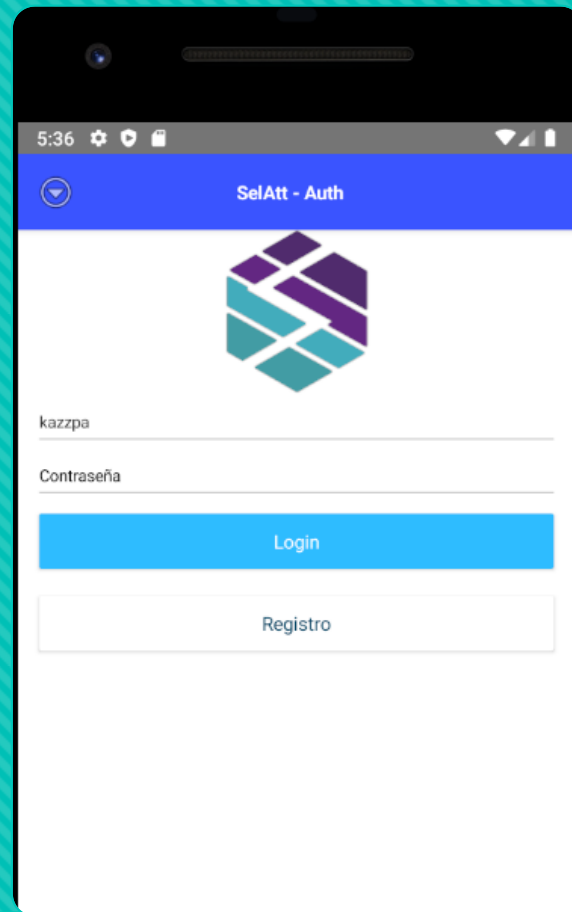
1. Objetivos

Sistema de información enfocado a analistas de datasets biomédicos.

Ejecutar algoritmos de machine learning sobre datos biomédicos.

Incorporar algoritmos de distintas fuentes.

Acceso multiplataforma.



2. Solución Propuesta

2.1 API PÚBLICA

- Autenticación
- Datasets
- Selección de atributos
- Clasificación
- Resultados
- Notificaciones

Autenticación

POST auth/login Devuelve un JWT de autenticación para las peticiones. ▾

POST auth/register Registro en el sistema. ▾

Datasets

GET featureSelection/datasets Devuelve los datasets subidos ▾

POST fileManager/uploadDataset Subir un dataset para procesarlo posteriormente. ▾

Selección de atributos

GET featureSelection/{algorithm}/{datasetName} Devuelve el resultado de realizar una selección de atributos ▾

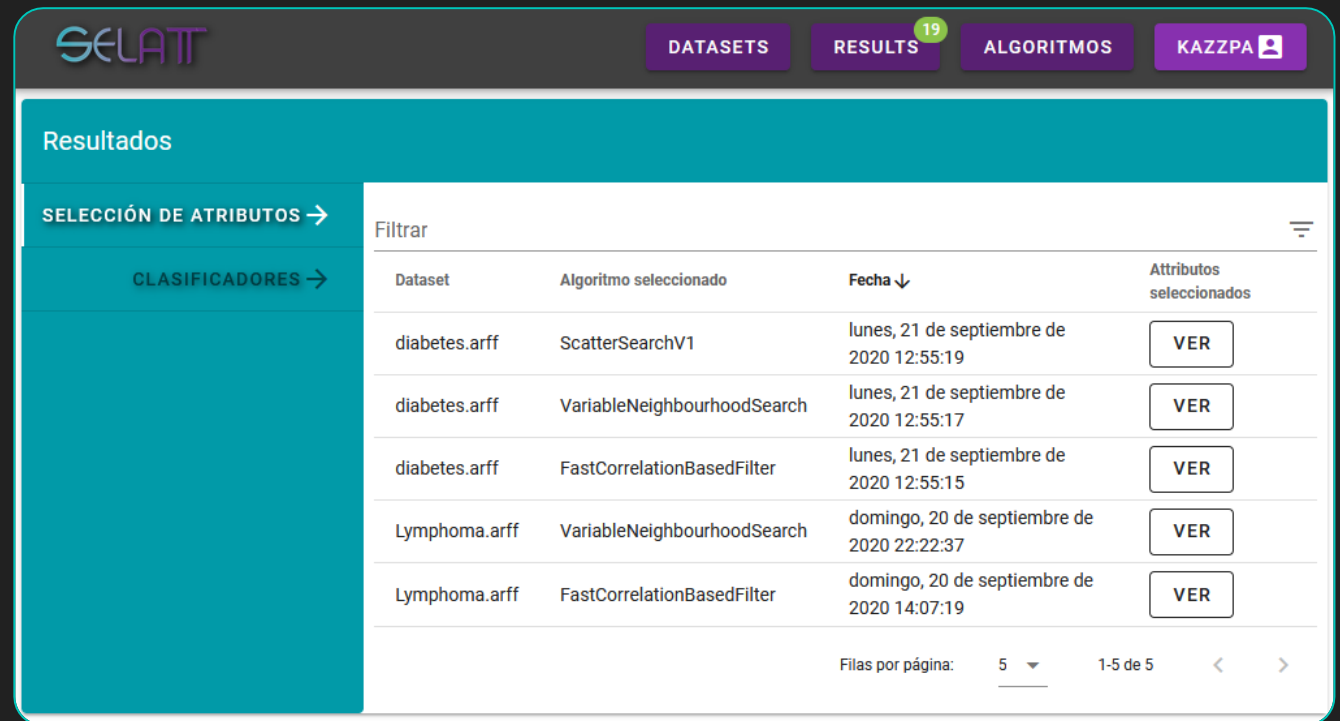
GET featureSelection/results Devuelve todos los resultados ▾

GET featureSelection/results/new Devuelve los resultados que no han sido vistos aún ▾

POST featureSelection/result/seen Marca el resultado como visto. ▾

2.2 Aplicación Web

- Acceso multiplataforma
- Análisis de datos
- Guía de uso API Pública

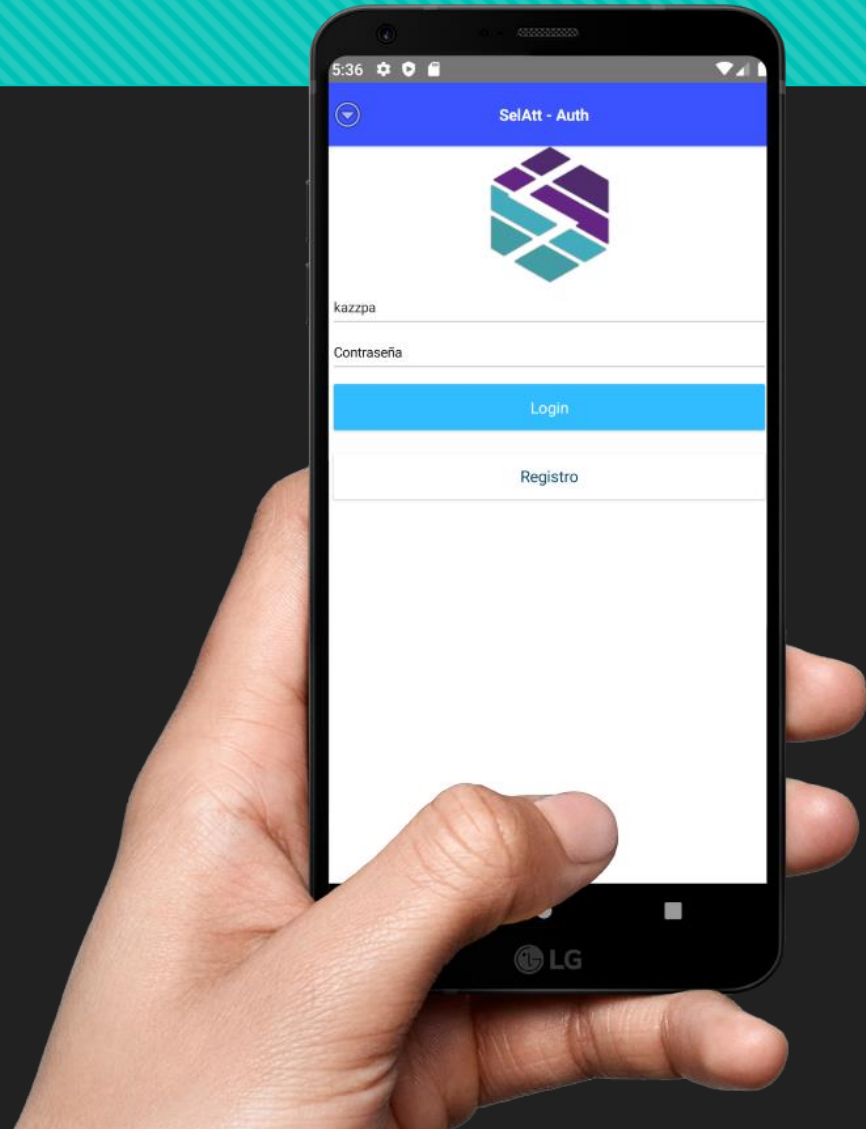


The screenshot displays the SELATT web application interface. The top navigation bar includes the SELATT logo and four buttons: DATASETS, RESULTS (with a green badge showing '19'), ALGORITMOS, and KAZZPA with a user icon. The main content area is titled 'Resultados' and features a sidebar with 'SELECCIÓN DE ATRIBUTOS →' and 'CLASIFICADORES →'. The main table, titled 'Filtrar', lists processing results for two datasets: diabetes.arff and Lymphoma.arff. Each dataset has two entries for different algorithms: ScatterSearchV1 and VariableNeighbourhoodSearch, and FastCorrelationBasedFilter. The table columns are Dataset, Algoritmo seleccionado, Fecha ↓, and Atributos seleccionados. Each row includes a 'VER' button. The bottom of the table shows pagination: 'Filas por página: 5' and '1-5 de 5'.

Dataset	Algoritmo seleccionado	Fecha ↓	Atributos seleccionados
diabetes.arff	ScatterSearchV1	lunes, 21 de septiembre de 2020 12:55:19	VER
diabetes.arff	VariableNeighbourhoodSearch	lunes, 21 de septiembre de 2020 12:55:17	VER
diabetes.arff	FastCorrelationBasedFilter	lunes, 21 de septiembre de 2020 12:55:15	VER
Lymphoma.arff	VariableNeighbourhoodSearch	domingo, 20 de septiembre de 2020 22:22:37	VER
Lymphoma.arff	FastCorrelationBasedFilter	domingo, 20 de septiembre de 2020 14:07:19	VER

2.3 Aplicación móvil nativa

- Accede a resultados desde cualquier lugar.
- Envía peticiones de nuevos algoritmos



2.4 Algoritmos disponibles

Selección de atributos:

Externo:

- Variable Neighbourhood Search

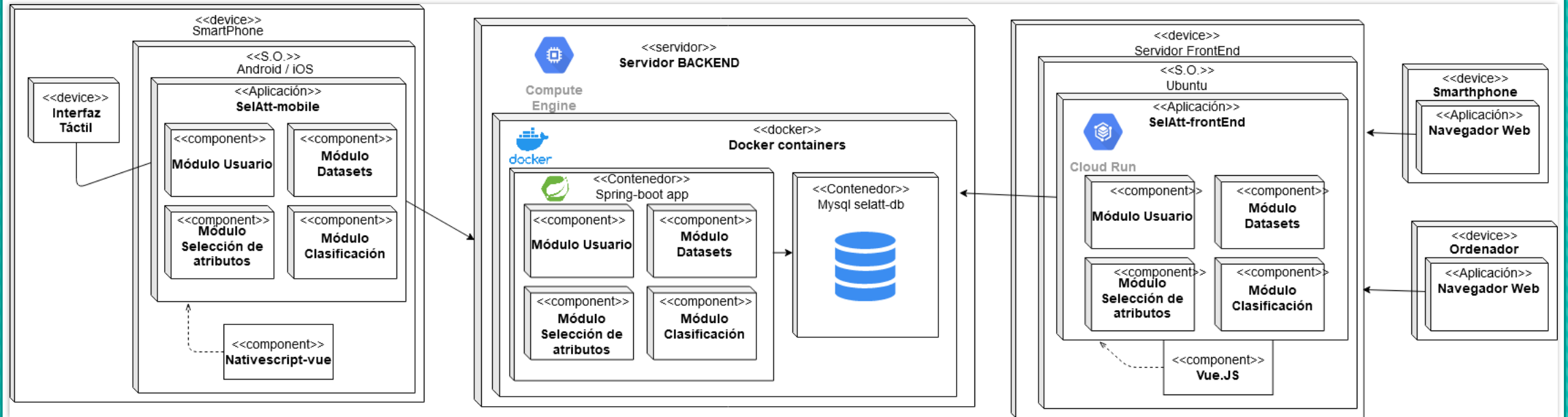
Weka:

- Fast Correlation Based Filter
- Ranker
- Scatter Search
- Exhaustive Search

Clasificadores:

Weka:

- Naive Bayes
- Naive Bayes Net
- Hidden Naive bayes
- Ibk
- Multi Layer Perceptron



3. Arquitectura

4. Tecnologías

Servidor BackEnd



Weka

Aplicación Web



Aplicación Móvil



NativeScript



Nativescript-Vue

5. Demo



Projet Nélanine

pourcentage de jeunes directrices au Bénin?

équilibre social?

catégories d'âge?





Machine Learning aplicado a datos Biomedicos

Aplicación de data mining especializada en selección de atributos a servicio mediante frameworks de última generación como Spring Boot y Vue.js

Las cookies se utilizan para mejorar la experiencia de uso de la aplicación, pero no son necesarias.

ACEPTAR

CANCELAR

6. Experimentación

- Reducción de atributos
- Solucionamos problemas asociados a datos biomédicos
- Herramienta para investigadores
- Lista de atributos descargable en app web.

Diabetes	Sin filtrar	VNS	FCBF
Naive Bayes	75.52%	68.49%	66.15%
Bayes Net	76.43%	68.49%	65.76%
Ibk	68.36%	60.42%	60.55%

Linfoma	Sin filtrar	VNS	FCBF
Naive Bayes	92.42%	92.42%	92.42%
Bayes Net	98.48%	92.42%	92.42%
Ibk	98.48%	98.8%	96.67%

Diabetes	Num atributos	Porcentaje
Sin Filtrar	8	100
FCBF	4	50
VNS	3	37,5

Linfoma	Num atributos	Porcentaje
Sin Filtrar	4026	100%
FCBF	110	2,73
VNS	62	1,53

7. Conclusiones

Objetivos cumplidos:

- Solución multiplataforma
- Competencias adquiridas
- Conocimientos del grado utilizados
- Aprendizaje desarrollo aplicación móvil

Futuras mejoras:

1. Descarga datasets completos filtrados
2. Añadir funcionalidad a aplicación móvil
3. Incorporar algoritmos fácilmente
4. Incluir automatización de tests para desplegar nuevas versiones rápidamente.

7. Agradecimientos

Preguntas