

## **RESUMEN**

En la actualidad, en casi cualquier campo de trabajo, se necesita procesar datos de manera más efectiva e instantánea. Uno de los campos que en Ecuador es imperativo se mejore, es el control en invernaderos para el correcto tratamiento de cultivos, sobre todo de exportación. En el campus IASA I de la Universidad de las Fuerzas Armadas ESPE, se lleva a cabo un proyecto para precisamente mejorar el control de su invernadero de rosas, dentro de este proyecto se han buscado las mejores herramientas en el campo de flujos de datos continuos con Big data como parte de la respuesta. Se identificaron dos herramientas idóneas para el tratamiento de flujos de datos provenientes del invernadero que son Spark Streaming y Apache Flink, se realizó una comparativa basándose en la norma ISO 25000 y se obtuvo como resultado que la herramienta Spark Streaming fue la idónea para el procesamiento de estos datos. Los datos para la solución que brinda Spark streaming, se encuentran almacenados en archivos y fueron procesados de manera que se discriminaba datos correctos o incorrectos de acuerdo a la temperatura y hora que fue transmitida desde uno de los sensores dispuestos en el invernadero. Como resultado se encuentra que efectivamente la herramienta fue la indicada para los flujos de datos, y se anima a realizar más verificaciones y a interactuar con Spark streaming y otros de los sensores de las rosas, siempre con la primicia de mejorar el control del mismo.

## **PALABRAS CLAVES**

- **BIG DATA**
- **SPARK STREAMING**
- **APACHE FLINK**
- **INVERNADEROS**
- **FLUJOS DE DATOS CONTINUOS**

## **ABSTRACT**

Nowadays, almost in any field of work, it is needed to process data effectively and instantaneously. One of the imperative fields of improvement in Ecuador is the control of greenhouses for the correct treatment of crops, especially for export. On the IASA I campus of the University of the Armed Forces ESPE, a project is being carried out to precisely improve the control of its roses greenhouse, within this project the best tools in the field of data streaming with Big Data, have been searched as part of the response. Two suitable tools were identified for the treatment of datastream from the greenhouse, Spark Streaming and Apache Flink, a comparison was made based on the ISO 25000 standard and the result was that the Spark Streaming tool was the ideal tool for processing these dates. The data for the solution provided by spark streaming, are stored in files and were process in a way that we discriminate them like correct or incorrect according to the temperature and time that was transmitted from one of the sensors arranged in the greenhouse. As a result, we found that the tool was indeed the one indicated for this datastream, and we encourage the community to carry out more verifications and to interact with Spark streaming and other sensors of the roses, always with the scoop of improving the control.

## **KEY WORDS**

- **BIG DATA**
- **SPARK STREAMING**
- **APACHE FLINK**
- **GREENHOUSES**
- **DATA STREAMING**