


<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 1 de 52

**Modelo de Selección de Candidatos usando técnicas de Machine Learning y
Tecnología Blockchain**

Yesenia Deantonio Saenz

Josefina Arias Rojas


Facultad de Ingeniería, Diseño e Innovación

Proyecto de grado

Tutor: Wilson Eduardo Soto Forero


Bogotá D.C

Diciembre 2023

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 2 de 52

Contenido

1. RESUMÉN.....	3
Palabras Clave.....	3
3. ALCANCE.....	6
4. PLANTEAMIENTO DEL PROBLEMA DE INVESTIGACIÓN	7
5. OBJETIVOS DE INVESTIGACIÓN	9
Objetivo General.....	9
Objetivos Específicos	9
6. REVISIÓN DE LITERATURA.....	10
7. DISEÑO METODOLÓGICO.....	22
8. RESULTADOS	36
9. CONCLUSIONES Y RECOMENDACIONES.....	44
10. REFERENCIAS BIBLIOGRÁFICAS.....	46

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 3 de 52

1. RESUMEN


En consecuencia de la creciente necesidad en las empresas de fortalecer el área de recursos humanos para el proceso de reclutamiento, en el presente trabajo se propone el diseño de un modelo basado en machine learning este enfoque implicó la integración de algoritmos de agrupamiento y clasificación, con base en el nivel de conocimiento de las habilidades técnicas de los posibles candidatos, de manera que se logró realizar un análisis de resultados del proceso de agrupamiento que luego se tuvo encuentra para el modelo de clasificación del candidato con el nivel de habilidad respectivamente, este proceso es crucial para que exista una categorización de los posibles candidatos, facilitando la selección de personal a perfiles específicos de la empresa; una vez realizada la clasificación del candidato pre seleccionado y garantizando la integridad y transparencia de los datos, se implementó un contrato inteligente con basado en la tecnología blockchain como etapa final del proceso.

Palabras Clave

Reclutamiento, Machine learning, Algoritmo, Agrupamiento, Clasificación, Blockchain, Contrato Inteligente.

Keywords

Recruitment, Machine learning, Algorithm, Clustering, Classification, Blockchain, Smart Contract.


<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 4 de 52

2. INTRODUCCIÓN

Vivimos en una era de constante cambio, donde la tecnología desempeña un papel central en la innovación y el aprendizaje en diversos ámbitos, tanto profesionales como sociales y culturales. En esta transformación tecnológica, se destaca la importancia de la integración de nuevas tecnologías para impulsar el progreso en las organizaciones. Esta integración se vuelve especialmente crucial en un entorno empresarial en constante evolución, donde la necesidad de adaptación es relevante.

Un área donde la tecnología ha traído cambios profundos en el proceso de reclutamiento y selección de personal. La manera en que las empresas atraen, identifican y seleccionan talento humano se ha transformado gracias a las nuevas tecnologías, evolucionando la manera en que las empresas generan vínculos para la contratación, mediante las redes sociales y estrategias de visualización para mejorar la manera en cómo es vista una organización y que tan atractiva puede llegar a ser para nuevos candidatos, además de implementar estrategias basadas en la búsqueda de nuevos talentos teniendo en cuenta el perfil profesional de los posibles candidatos convirtiéndose en una práctica esencial.

El área de reclutamiento, que representa uno de los pilares críticos de la gestión de recursos humanos en las organizaciones, se concentra en el filtrado, la selección y la contratación de candidatos para nuevas vacantes, de manera que sea posible la identificación del valor agregado que cada individuo puede aportar a la organización, adicional se encuentra en la constante


<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 5 de 52

búsqueda de herramientas para optimizar procesos y analizar información relevante para las organizaciones y aportar a la gestión de recursos humanos, la inteligencia artificial ofrece beneficios para aprender, entender y tomar decisiones basada en datos, permitiendo el desarrollo del machine learning, que atribuye la participación de algoritmos como procesos que revolucionan la manera en que las empresas identifican y seleccionan nuevos talentos, además de la reducción de sesgos en la selección y clasificación de las características requeridas para un cargo.

El sesgo en la selección de personal es un problema común, donde los prejuicios o estereotipos pueden influir en las decisiones de contratación, a menudo de manera inconsciente. Los algoritmos basados en inteligencia artificial se han convertido en una solución eficaz a este desafío. Estos algoritmos permiten la evaluación objetiva, sin influencias personales, con el objetivo de reducir significativamente el sesgo en la selección de candidatos y mejorar la toma de decisiones.


En el presente trabajo, se busca exponer el uso de la inteligencia artificial como una herramienta capaz de realizar tareas y tomar decisiones basadas en razonamiento y datos. La tecnología de machine learning y el blockchain se utilizan para analizar, procesar, almacenar y demostrar el autoaprendizaje mediante la implementación de algoritmos de agrupamiento y clasificación, permitiendo optimizar el proceso de reclutamiento masivo en las organizaciones.

Este enfoque no solo agiliza la selección de candidatos, sino que también contribuye al crecimiento de las organizaciones, con personal que genera valor y contribuyen al éxito de las empresas.

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 6 de 52

3. ALCANCE


Diseñar un algoritmo de agrupamiento y clasificación con base en un conjunto de datos aleatorios que posee un candidato que se postula a una vacante, haciendo uso de la herramienta de Knime para implementar un flujo de trabajo compuesto por los nodos de cada algoritmo, la data que se requiere evaluar y la métrica de evaluación de asertividad y eficiencia de cada algoritmo, adicional teniendo en cuenta los resultados, se implementa un contrato inteligente en lenguaje solidity, cuyo despliegue se realiza en una blockchain local.

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 7 de 52


4. PLANTEAMIENTO DEL PROBLEMA DE INVESTIGACIÓN

En un entorno empresarial y altamente competitivo, existen variedad de desafíos a los que se enfrentan las grandes empresas en el área de recursos humanos, donde se refleja la falta de recursos y personal, la ausencia de tecnologías para optimizar procesos, falta de visión estratégica, dificultades con el procesamiento de información, entre otros, que a la final se ve reflejados en los resultados del crecimiento de la organización, de manera que se busca implementar nuevas tecnologías que permiten dar solución a estos problemas, en temas de rendimiento, mejora operativa, innovación y eficiencia operativa.

En términos del reclutamiento masivo, son muchas las empresas que pueden manejar entre 30 y 40 vacantes en un mismo proceso, en las cuales se pueden presentar hasta 250 candidatos lo que le representa a un profesional de recursos humanos, visualizar, clasificar y seleccionar entre una gran cantidad de postulaciones al año y si además carece de las herramientas adecuadas para el proceso de reclutamiento y clasificación de personal puede ser un verdadero desafío. Adicionalmente, desde el punto de vista de los candidatos y gracias a las plataformas digitales que son el puente de contratación entre empresas y candidatos, se pueden recibir variedad de postulaciones de candidatos que no cumplen con las características especificadas en el puesto y puede ser difícil de filtrar la información que se requiere para iniciar con el proceso de reclutamiento y contratación, sumando los esfuerzos adicionales que se requieren para capacitar a un individuo una vez es contratado si no cumple con las características solicitadas. En este contexto surge la necesidad de mejorar la eficiencia y confiabilidad en los procesos de reclutamiento implementado tecnologías que aporten a la mejora de estos procesos.

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 8 de 52

Apoyarse en las nuevas tecnologías como Blockchain para el almacenamiento seguro de información y la inteligencia artificial para el agrupamiento, clasificación y refinamiento de información, permite contribuir a la solución de los desafíos a los que se enfrenta el área de recursos humanos, de manera que se pueda ver una reducción de tiempo y costos, contribuyendo a la búsqueda de nuevos talentos y una selección exitosa.

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 9 de 52


5. OBJETIVOS DE INVESTIGACIÓN

Objetivo General

Realizar un modelo de machine learning y Blockchain que apoye el proceso de selección de candidatos en una empresa.

Objetivos Específicos

- Diseñar un algoritmo de técnica no supervisada con un conjunto de datos históricos de candidatos con niveles en habilidades técnicas.
- Analizar los resultados obtenidos con el algoritmo de técnica no supervisada.
- Diseñar un algoritmo de técnica supervisada usando los datos agrupados resultantes del algoritmo de técnica no supervisada.
- Construir un contrato inteligente basado en la tecnología del blockchain para el candidato seleccionado.


<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 10 de 52

6. REVISIÓN DE LITERATURA

A inicios del siglo XX en consecuencia de la Revolución industrial surge el primer concepto para el área de recursos humanos llamado Relaciones industriales, cuya finalidad consistía en mediar entre las organizaciones y las personas, de manera que fuera posible enlazar el capital y el trabajo (Chiavenato A. , 2011), en relación al surgimiento de esta área se fortalece la importancia del recurso humano para las diferentes organizaciones y el éxito de las mismas, sin embargo hasta después de los 90 se empieza a considerar a las personas como miembros que generan valor, impulsan la productividad y el aprendizaje continuo, para alcanzar los objetivos y metas de las organizaciones.


Desde el punto de vista de los individuos que componen el área de recursos humanos cuando se habla de una organización con años de trayectoria, que adicional maneja procesos grandes según las diferentes actividades en las que se desempeña, este tipo de empresas requieren realizar contrataciones masivas de personal, que sin los recursos necesarios podría ser un problema para el área de recursos humanos y en general afectar la organización directa o indirectamente en el punto de la productividad organizacional. El reclutamiento masivo es un proceso de selección de personal requerido para un mismo puesto u otros en específico, con el objetivo de cubrir procesos grandes y/o vacantes antiguas (Guerrero Barrera, 2017).

Actualmente se presentan varias maneras de generar empleo a través de diferentes plataformas gracias a internet y redes sociales, con la posibilidad de hallar diferentes cargos de acuerdo al perfil del individuo y lograr postulaciones de manera más fácil y rápida desde cualquier parte del mundo,

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 11 de 52


a este proceso se le llama reclutamiento 3.0; teniendo en cuenta que así como aumentan la cantidad de postulaciones a los diferentes cargos de las empresas, también se refleja un aumento en la selección de candidatos y el filtro que se debe llevar a cabo para lograr una contratación, basándose en las competencias requeridas para un cargo, que servirá para determinar los criterios para seleccionar y descartar candidatos.

El proceso de reclutamiento ha mantenido una evolución constante hasta el denominado reclutamiento 4.0, con enfoque en el manejo e innovación de las TICS y la implementación de tecnologías para llevar a cabo este proceso de manera correcta y eficiente, adicional enfoca sus esfuerzos en la implementación de técnicas de marketing digital, como lo es el employer branding como un concepto para realzar la imagen de la empresa generando posicionamiento y reputación con el objetivo de proyectar y atraer nuevos talentos. En la industria 4.0 sobresalen algunas tecnologías habilitadoras que impulsan la digitación en la industria, se encuentra la inteligencia artificial que responde a un programa de computación capaz razonar y presentar capacidades similares a un ser humano, la cual mantiene una lógica desarrollada, para el procesamiento de información y la toma de decisiones a través de algoritmos y mecanismos que permiten la negociación automatizada para diferentes escenarios y criterios planteados para esa toma de decisiones (Aydoğan, Marsa-Maestre, Klein, & M. Jonker , 2018), la manera para aplicar estos algoritmos es mediante el Machine Learning o aprendizaje automático, elemento relevante de la ciencia de datos.

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 12 de 52

El machine learning como rama que hace parte de la inteligencia artificial, maneja variedad de algoritmos con la capacidad de aprender buscando patrones, sin necesidad de ser programados, esto se logra a través de un gran volumen de datos permitiendo que un algoritmo seleccionado aprenda y sepa cómo comportarse en cada situación requerida con un porcentaje mínimo de error (Sandoval, 2018). A diario se encuentran diferentes aplicaciones donde es aplicada la técnica del machine learning como el filtrado de correos, el reconocimiento de voz, la predicción del clima y la detección del tráfico, todo a través de patrones que son encontrados (Maisueche-Cuadrado, 2019).


Cuando se realiza un modelo basado en el machine learning existen algunos puntos para esta implementación, teniendo en cuenta el porcentaje de error que esta puede llegar a tener, el sesgo y el sobreajuste que se puede requerir. El aprendizaje se basa en el planteamiento de parámetros en función de datos iniciales (data set) presentados, que serán analizados por el algoritmo seleccionado, obteniendo un resultado inicial esperado, es importante tener en cuenta que se debe seleccionar un algoritmo que se ajuste a las necesidades de cada modelo para obtener resultados más acertados y para poder implementar cada algoritmo se tiene en cuenta que el conjunto de data set, se divide en dos subconjuntos, donde el 80% de estos datos son utilizados para entrenar (training data) al algoritmo con los datos proporcionados y el subconjunto de datos de prueba (test data) que corresponde al 20% de los datos para obtener la información esperada después de entrenar al algoritmo (Maisueche-Cuadrado, 2019).

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 13 de 52

La técnica del machine learning aplicada a una empresa, específicamente en el área de recursos humanos para analizar grandes volúmenes de información, teniendo en cuenta la selección de características específicas basadas en el perfil de diferentes candidatos que aplican a una vacante, para ello es posible manejar un método de clustering para identificar grupos o categorías de un conjunto de datos, este se basa en la agrupación de elementos con características similares, de esta manera creando un clúster (Bojorque-Chasi, 2020). Para realizar un proceso de agrupamiento de características definidas para una vacante, es importante reconocer que se debe trabajar con algoritmos de aprendizaje no supervisado, ya que es una técnica cuya finalidad es procesar la información a partir de un conjunto de datos y encontrar patrones o estructuras, que se encuentran ocultos en los datos, haciendo que estos hablen por si solos (Diaz-Romero & Sanyer Mosquera, 2021).

En el caso del objetivo del área de reclutamiento de una empresa, es necesario hacer uso del algoritmo no supervisado bajo el método de clúster o agrupamientos, en este caso se estudia sobre el algoritmo K-Means creado por MacQueen en el año 1967, cuyo objetivo se centra en la separación de un conjunto de elementos en una cantidad K de clusters, que se representan como centroides y se ubican arbitrariamente al momento de procesar el algoritmo, el nombre del algoritmo se origina a partir de los K grupos que pueden formarse con base a la distancia euclidiana hay existe entre los objetos y los centroides; se desarrolla en 4 etapas:

Etapas 1: Elección de K objetos de manera aleatoria, para K clusters iniciales


<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 14 de 52

Etapa 2: Asignación de objetos a los K clusters, cada objeto debe ser asignado al cluster más cercano

Etapa 3: En cuanto se ubican los objetos, se realiza una nueva iteración de los K cluster

Etapa 4: Se deben repetir las etapas 2 y 3 hasta que todos los objetos se encuentren con el cluster más cercano, es decir no hallan más reasignaciones (García-Cambronero & Gomez-Moreno, 2006)

El algoritmo K-means siempre finaliza la ejecución, pero no hay garantía que la solución final sea la más óptima, ya que esto depende de varios factores como los datos proporcionados, las iteraciones para el acercamiento de clusters y formación de agrupaciones, adicional es importante mencionar que el algoritmo K-Means proporciona ahorro de costos y mejora de rendimientos (Zahra, y otros, 2015). Por otro lado, se encuentra un siguiente algoritmo que busca para generar agrupaciones a través de clusters llamado algoritmo jerárquico, el cual particiona un conjunto de datos por niveles, de manera que se unen dos objetos con la menor distancia creando grupos, para luego dos grupos unirse formado un grupo más grande (Pascual, Pla, & Sánchez, 2007), este es un método en el que se pueden agrupar los datos llamado aglomerativo, sin embargo existe otro método donde se inicia con un solo grupo y va separando los datos en clusters, este método se refiere al divisivo (Monje López, 2021), ambos métodos se visualizan en la figura 1

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 15 de 52

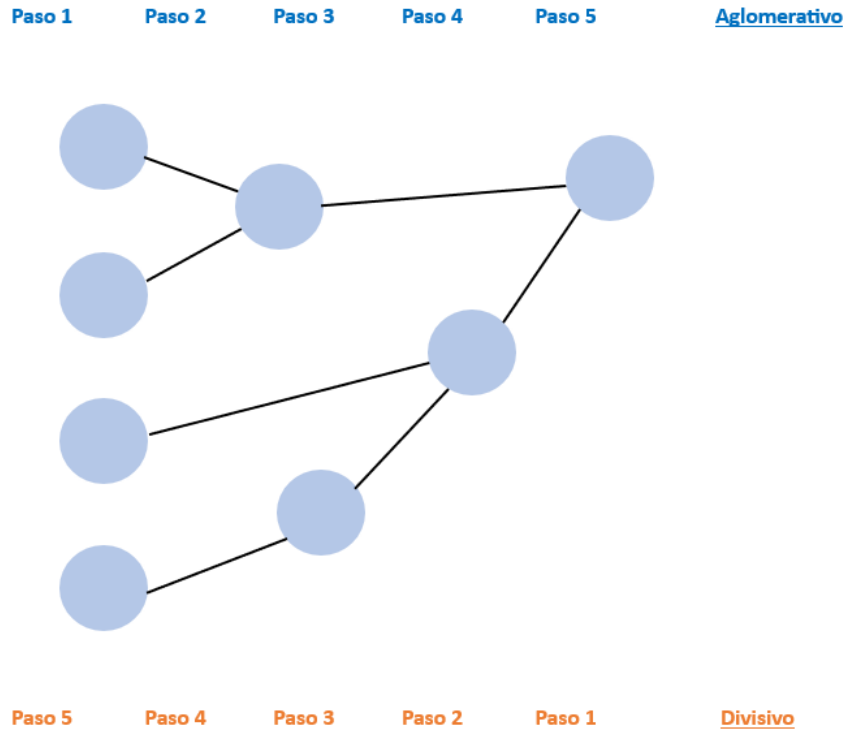



Figura 1: Métodos de agrupamiento, Imagen Propia, 2023 Fuente. Microsoft Word

Adicional, el algoritmo presenta una jerarquía de conglomerados que pueden ser visualizados en un dendograma donde en el eje X se visualiza el número de clusters y el eje Y hace referencia a la distancia como se muestra en la figura 2, de modo que se pueden presenciar los posibles clusters a partir de los grupos formados (Rebafka, 2023).

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 16 de 52

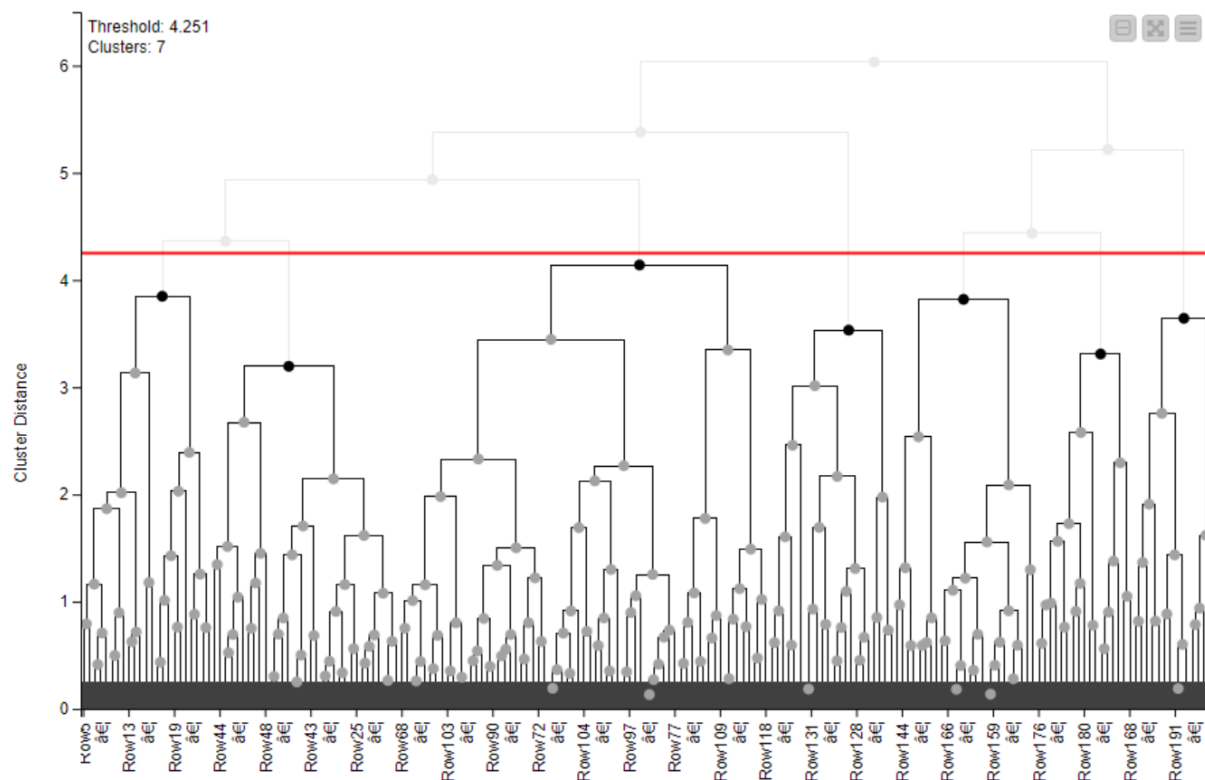



Figura 2: Dendrograma, Imagen Propia, 2023 Fuente. Knime Analytics Platform


De acuerdo con la figura anterior, el dendrograma permite definir que cantidad de cluster es mas conveniente tomar, según el objetivo individual de la agrupación.

Para procesos de reclutamiento también es utilizado el algoritmo de árbol de decisión, seleccionado como de los más sencillos de implementar, buscando la generación de un árbol recursivo, representado por un conjunto de nodos, hojas y ramas, los nodos se dividen entre el nodo principal o raíz que es donde se inicia la creación del árbol y nodos internos que abarca a los nodos hijos, nodos finales (hojas) y las ramas que representando los posibles caminos de un nodo a otro (Barrientos-Martinez, y otros, 2009), los árboles de decisión son una técnica considerada en

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 17 de 52

la minería de datos, principalmente para el proceso de predicción, clasificación y segmentación. Para la implementación de este algoritmo en un flujo de trabajo se induce un conjunto de valores de atributos, conocidos como datos de entrenamiento y la definición de unas clases objetivo que se refiere a la categorización a la que va orientada los resultados. En primera instancia se selecciona el atributo que mejor divide el conjunto de datos, para que sea el nodo principal, los demás nodos se dividen en subconjuntos, de acuerdo con el primer atributo del paso 1 para conformar los nodos hijos, estos dos pasos se repiten por cada subconjunto de datos hasta llegar a los nodos hojas que son lo que finalizan una ramificación de nodos, teniendo en cuenta algún requerimiento de parada especificado (Su & Zhang, 2006).

Cada algoritmo cuenta con métricas para evaluar la precisión y alto porcentaje de rendimiento, por ejemplo, para K-Means y los árboles de decisión, es posible hacer uso de una herramienta específicamente para modelos de clasificación, en este caso se habla de la matriz de confusión la cual contiene métricas para identificar la precisión del modelo, esta herramienta permite realizar una comparación de casos positivos y negativos basados en el modelo, donde se toma un porcentaje de casos de prueba para entrenar el algoritmo con el conjunto de datos iniciales y el resto de los casos son los reales para dar un resultado más preciso, la matriz divide los resultados en 4, los verdaderos positivos, haciendo referencia a los casos que se predijeron correctamente, a los falsos negativos, siendo la cantidad de casos que se predijeron como negativos, pero en realidad son positivos, los falsos positivos siendo la cantidad de casos que se predijeron como positivos, pero en realidad son negativos y por último los verdaderos negativos, siendo la cantidad de casos que

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 18 de 52


el modelo predijo como negativos y es correcto (Diaz-Romero & Sanyer Mosquera, 2021), la matriz de confusión se visualiza en la figura 3.

		1	0
Predicción de valores	1	TP	FP
	0	FN	TN

Figura 3: Matriz de confusión, Imagen Propia, 2023 Fuente. Microsoft Excel

Para el algoritmo jerárquico es posible hacer uso de la métrica del coeficiente de silueta, ya que permite comparar y calcular la cercanía de los clusters y su acertada agrupación, de manera que se mide en un rango de $-1 < s(i) < 1$, donde $s(i) = 1$, el dato i está asignado correctamente, $s(i) = 0$, el dato i se encuentra entre dos posiciones y $s(i) = -1$, el dato i se encuentra mal asignado, de esta manera es factible identificar la correcta agrupación del algoritmo (Jacanamejoy-Jamioy, Forero-Vargas, & Tavera-Bucurú, 2019)

En término de procesamiento de información, se selecciona Knime, una plataforma de código abierto con interfaz gráfica para la implementación del Machine Learning, Business intelligence y ETL que por sus siglas en inglés hace referencia a Extract para la obtención de datos, Transform que adapta los datos y Load haciendo referencia al proceso de cargue de la información (Diaz-Romero & Sanyer Mosquera, 2021), Knime permite la creación de flujos de trabajo (Workflows) conformados por un conjunto de nodos definidos para un fin específico, cada nodo posee cuatro diferentes estados:

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 19 de 52

Nodo Rojo: Nodo sin configurar


Nodo Amarillo: Nodo sin ejecutar ya configurado

Nodo Verde: Nodo ejecutado

Nodo Rojo con cruz blanca: Nodo ejecutado con falla (García-Bermejo, 2022)

Knime es una herramienta que permite visualizar un proceso de clasificación de manera más eficiente y sencilla (Sapriani, Iwan, Windo, & Erlinawaty, 2021), ya que cuenta con elementos para facilitar su uso como lo son los nodos que cumplen con una función específica, desde lectura de archivos externos hasta predicciones de modelos, todos los nodos al ser implementados y conectados entre sí, se le da el nombre de workflow siendo la secuencia de los mismos nodos. Todos los nodos que maneja Knime, son categorizados y subcategorizados de manera que sea más fácil de encontrar; esta herramienta cuenta con nodos IO (Input – Output), con subcategorías Read (leer) y Write (escritura), también cuenta con nodos DB (base de datos) y nodos Query para realizar consultas específicas con la información proporcionada, nodos de control de flujos, Switches para implementar bucles, condicionales y tratamiento de errores, nodos para la obtención de mejores visualizaciones y para realizar operaciones y análisis, como también nodos con la implementación de diferentes algoritmos de clasificación, agrupación, regresión entre otros (Emilio, 2020).


Una vez analizadas herramientas para el procesamiento de datos y continuando con el proceso de reclutamiento en una empresa, es preciso tener en cuenta no solo el proceso de clasificación mediante técnicas de machine learning si no también buscar la manera de realizar este proceso de manera más eficaz, donde después de una clasificación y selección, exista un procedimiento de

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 20 de 52

creación de contratos para hacer más efectiva la contratación, para ello se habla del blockchain, una tecnología de registro de información almacenada.

En el 2017 inicia el auge de la tecnología blockchain con la valorización del bitcoin, cuya creación se atribuye a Satoshi Nakatomo quien en el 2008 publica un documento basado en el doble gasto, donde enfatiza en la moneda digital la cual puede ser duplicada y ser utilizada más de una vez, dicho problema fue resuelto a partir de vinculación de una transacción a partir de la anterior, confirmando que la moneda no ha sido gastada aún, es decir que cada vez que se genere una transacción entre dos nodos, la transacción es almacenada en un bloque con capacidad limitada, mientras que los nodos de la misma red trabajan en la verificación y confirmación de la transacción, teniendo en cuenta lo anterior surge el concepto del blockchain (LegalTech news, 2018). El Bitcoin hasta ahora ha sido la moneda con mayor éxito a partir del uso del blockchain, de esta manera pautando el inicio de las criptomonedas. (Bartoletti & Pompianu, 2017)


La tecnología del Blockchain se ha implementado para la creación de monedas descentralizadas, como se mencionó en el párrafo anterior, el Bitcoin es un ejemplo de ello, adicional para la creación de SmartContracts y Smartproperty (IEEE, 2016). Según (Golosova & Romanovs, 2018) “el Blockchain es un libro de contabilidad digital incorruptible de transacciones económicas que se puede programar para registrar no solo transacciones financieras, sino prácticamente todo lo que tenga valor”; esta declaración es una de las definiciones más populares de Blockchain, desarrollada por Don y Alex Tapscot. Esta tecnología se identifica por ser una cadena de bloques con datos distribuidos o una forma de registrar información a través del criptoanálisis, donde cada

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 21 de 52

bloque almacena una cantidad de registros y la relación del bloque anterior, por lo que es inamovible, es decir cada bloque que conforma la cadena se mantiene fijo. (Tasende, 2020)

Pero las características que hacen que esta tecnología sea tan atractiva para las nuevas implementaciones son descentralización, inmutabilidad, escalabilidad y privacidad limitada. El objetivo de la descentralización es descartar un rol intermediario entre las transacciones en una cadena de bloques, permitiendo que la información sea descentralizada y distribuida, de esta manera se garantiza sistemas y procesos transparentes y con la capacidad de generar confianza. (Gonzalez-Perez, 2018)

En el proceso de finalización de la transferencia de información en la cadena de bloques, se otorga aprobación, evitando que los bloques sean alterados. Esto garantiza la integridad de las transacciones registradas, especialmente en el caso de los Smart Contracts, donde la inmutabilidad se vincula con la ejecución automática al cumplir las condiciones predefinidas (Tasende, 2020). Los Smart Contracts son considerados como un programa de información que asegura y ejecuta acuerdos entre las partes interesadas, estos ayudan en el proceso de negociación de acuerdos y son programados para ejecutar tales términos definidos, sin la intervención humana (Valencia-Ramirez, 2019).

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 22 de 52

7. DISEÑO METODOLÓGICO

La Construcción del diseño metodológico se realizó en base a los objetivos definiendo tareas semanales, detallando el desarrollo de las actividades, así como sus dificultades, las cuales se dividen de la siguiente manera:

Semana 1: Recopilación de Datos

Tarea 1.1: Generar datos de los candidatos y sus habilidades técnicas de forma aleatoria.


Para iniciar con el análisis de la investigación, para el modelo de agrupamiento se realizó una generación aleatoria de 300 datos en los cuales se encuentran posibles candidatos, cada columna tiene su propio concepto, Backend, Front end, base de datos e infraestructura, denominados conocimientos y/o habilidades, discriminados con una puntuación de 1 a 100, donde 1 es que no tiene conocimiento y 100 tiene el conocimiento requerido para el cargo

Semana 2 - 3: Preparación de Datos

Tarea 1.2: Realizar un proceso de alistamiento de datos, incluyendo normalización y codificación de características relevantes.

En este paso para cada algoritmo se procesaron los datos en KNIME con los siguientes nodos:

Excel Reader: Este nodo implementado en el algoritmo de agrupamiento y clasificación lee los archivos de Excel en los diferentes formatos convirtiéndolo a los tipos de dato knime escaneando el archivo de entrada determinando el número y los tipos de columnas generando una tabla con la estructura adivinada automáticamente y los tipos KNIME.

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 23 de 52

Duplicate Row Filter: El Nodo implementado en el algoritmo de clasificación identifica las filas duplicadas que tienen valores idénticos en determinadas columnas del archivo leído, eliminando todas las filas duplicadas de la tabla de entrada.


Missing Value: El nodo implementado en el algoritmo de clasificación ayuda a manejar los valores faltantes en las celdas que encuentra de la tabla de entrada, permite configuraciones individuales para todas las columnas generadas.

Normalizer: Este nodo implementado en el algoritmo de agrupamiento normaliza los datos numéricos de todas las columnas con dicho tipo de dato, permite normalizaciones min-max, puntuación Z (Gaussiana), por escala decimal.

Semana 4 - 6: Implementación del algoritmo de agrupamiento

Tarea 1.3: Realizar la búsqueda e implementación de un algoritmo de agrupamiento, conociendo las características y configuraciones de cada nodo.

En el procedimiento de agrupación basado en la identificación de grupos con registros similares de acuerdo con las habilidades técnicas se buscaron algoritmos de agrupamiento que señalaran un buen rendimiento en la formación de grupos con el fin de implementar y asegurar la calidad de los grupos generados, Como podemos ver en la Figura 4, trabajamos con el algoritmo K-means implementando los siguientes nodos:

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 24 de 52

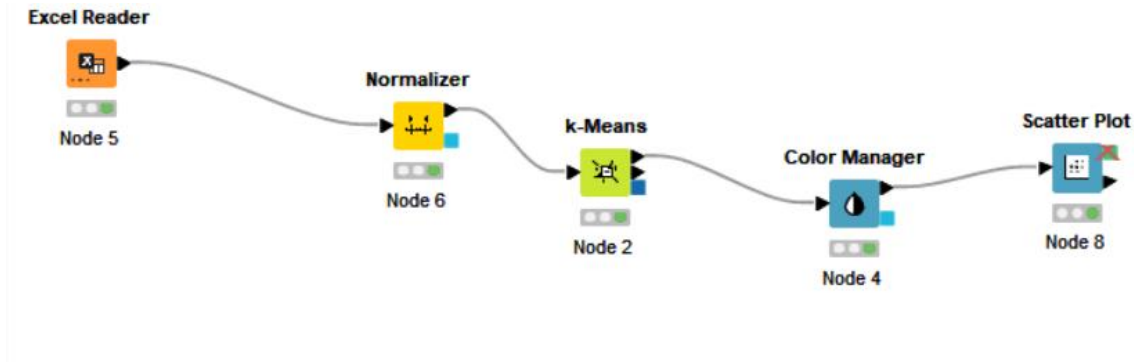



Figura 4. Modelo de Agrupamiento Kmeans, Imagen Propia, 2023 Fuente. Knime Analytics Platform

K-means: Genera unos centroides para el número de clústeres preestablecido realizando una agrupación nítida asignando un vector de datos a cada grupo, utiliza la distancia euclidiana en los atributos seleccionados.

Color Manager: Asigna colores a las columnas nominales con la etiqueta seleccionada, en este caso los clústeres definidos en el nodo anterior.

Scatter Plot: Este nodo como su nombre lo indica genera un gráfico de dispersión en las coordenadas(x,y), ayudando así a visualizar los grupos generados en cada centroide establecido.

Una vez generado el modelo de agrupamiento se es requerido analizar los grupos formados para identificar la pertenencia de cada uno, sin embargo, se descartó el algoritmo K-means ya que generando el grafico de dispersión los datos no se ubicaban en los centroides definidos y se visualizaban dispersos, de igual manera la información no se evidenciaba de forma correcta para ser analizada. Posterior a esto se realizó la búsqueda de un nuevo algoritmo y cumplir el objetivo

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 25 de 52

“Diseñar un algoritmo de técnica no supervisada con un conjunto de datos históricos de candidatos con niveles de habilidades técnicas”. Como podemos observar en la Figura 5, implementamos el algoritmo Jerárquico con los siguientes nodos:

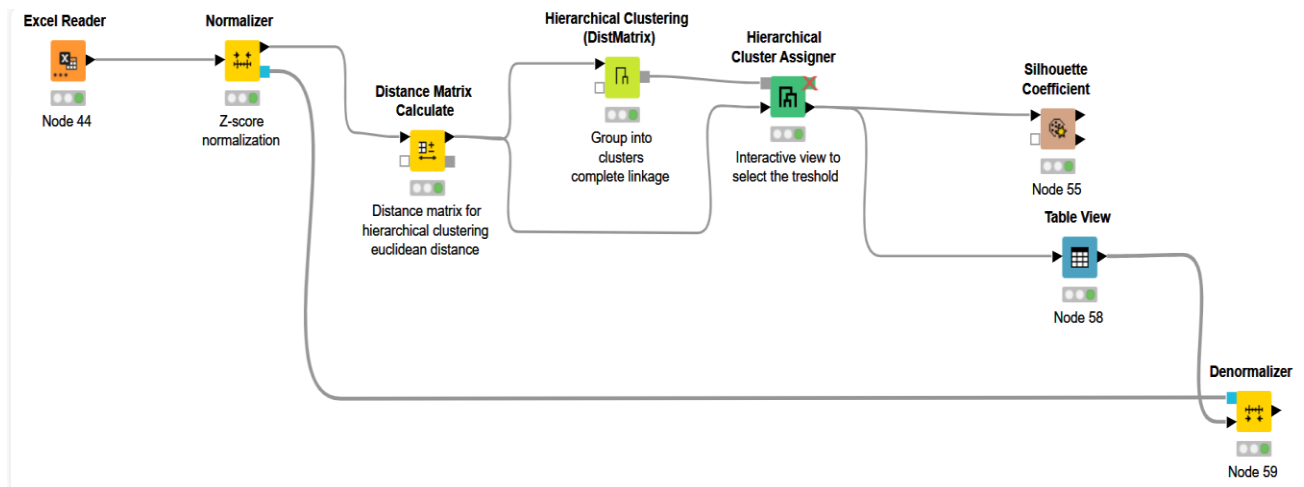



Figura 5. Modelo de Agrupamiento jerárquico, Imagen Propia, 2023 Fuente. Knime Analytics Platform

Distance Matrix Calculate: Este nodo calcula los valores de distancia para todos los pares de filas de la tabla generada en el nodo Excel Reader , como resultado agrega una columna con los valores del vector de distancia.

Hierarchical Clustering (DistMatrix): El nodo agrupa de forma jerárquica los datos entrantes mediante una matriz de distancia, para este fin existen dos métodos de agrupación jerárquica, el divisivo y el aglomerativo, el implementado funciona de forma aglomerativa lo que quiere decir que comienza con cada punto de datos como uno solo, intentando combinar los que tengan mayor similitud en subgrupos, hasta que finaliza con un gran grupo que contiene los subgrupos generados.

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 26 de 52

Para realizar el proceso mencionado se debe calcular la distancia entre los conglomerados definiendo una medida los métodos para realizar este proceso son enlace único, completo y promedio, para realizar un calculo con mayor precisión para el proceso se configuro el nodo con el enlace completo.

Hierarchical Clustering Assigner: El nodo tiene como entrada un árbol de cluster jerárquico y la tabla de entrada utilizada para la creación del cluster para posteriormente crear un dendograma interactivo con los clusters configurados, para la definición de los clusters se utiliza el nodo Silhouette Coefficient identificando la etiqueta overall que se encuentra entre -1 y 1 siendo los valores más cercanos a 1 el mejor puntaje, con esto se configuran nodos de 2 a 10 consignando el dato overall como se muestra en la Figura 6, en el grafico se evidencia que con 8 cluster el valor de overall es mayor, por lo anterior se identifica que con 8 clusters habrá mejor asignación y conformación de los grupos a generar.

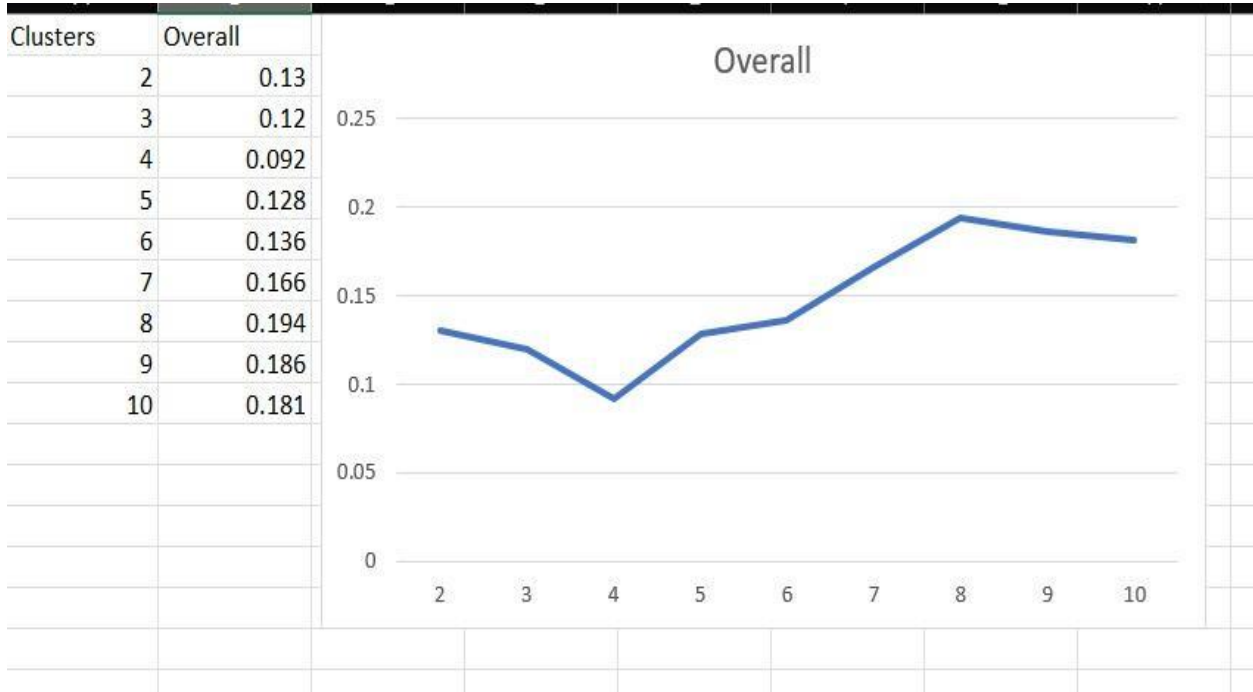



Figura 6. Análisis precisión Cantidad de Cluster, Imagen Propia, 2023 Fuente. Excel

Silhouette Coefficient: El nodo calcula el coeficiente de silueta para el resultado de la agrupación generada, es muy utilizado para evaluar el rendimiento, la puntuación del calculo que efectúa puede estar entre -1,0 y 1,0.

Table View: el nodo genera una tabla para visualizar los datos salientes del agrupamiento generado con el fin de obtener el cluster asignado a cada registro.

Desnormalizer: El nodo desnormaliza datos que anteriormente pasaron por el nodo normalizer, se implementa en el modelo con el fin de analizar los grupos generados y mejorar la lectura de los datos proporcionados en el nodo table view.

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 28 de 52

Semana 7 - 8: Análisis de Datos

Tarea 1.4 Analizar los datos generados en la tarea anterior, estando agrupados y con los cluster asignados a cada registro verificando las características principales en cada grupo.

Para efectuar el análisis de datos se procede a descargar el archivo generado en la tarea anterior resultado del nodo desnormalizer, en el cual podemos evidenciar los datos iniciales de las habilidades con un cluster asignado, se procede a generar la suma de cada habilidad por cluster a través de la tabla dinámica visualizada en la Figura 7 identificando la habilidad técnica que más se repetía en cada cluster, como se evidencia en la Figura 7.1, se generan grupos de a 2 cluster con características similares dejando así los grupos de las habilidades técnicas Infraestructura, Backend, FrontEnd y Base de Datos

Etiquetas de fila	Infraestructura Backend	FrontEnd	Base de Datos
Cluster_0	319	1496	1187
Cluster_1	1294	1569	1170
Cluster_2	2421	1854	1894
Cluster_3	1394	947	428
Cluster_4	1738	1444	1607
Cluster_5	956	865	2027
Cluster_6	1364	324	1187
Cluster_7	554	241	515
Total general	10040	8740	10015

Figura 7. Tabla Dinamica, Imagen Propia, 2023 Fuente. Excel

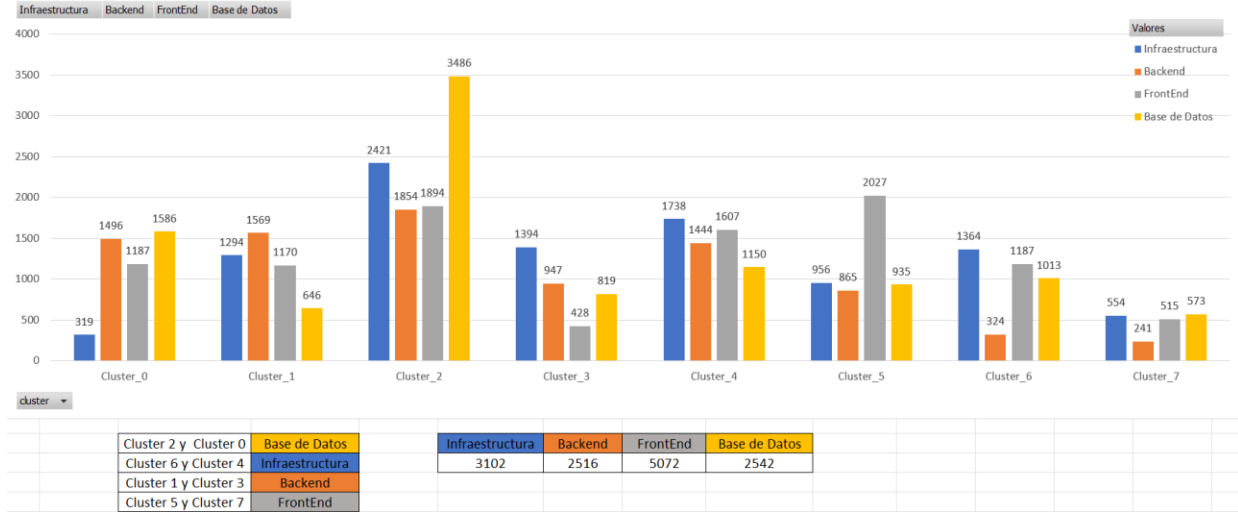



Figura 7,1. Análisis de Datos, Imagen Propia, 2023 Fuente. Excel

Semana 9 – 10: Implementación del algoritmo de clasificación

Tarea 1.5: Realizar la implementación de un algoritmo de clasificación

Para permitir la construcción del modelo predictivo requerido con algunas características configuradas, en este modelo de clasificación se busca predecir el valor del cluster o grupo en el que quedará cada candidato, es decir, se quiere pronosticar en que grupo o área quedara un candidato con las habilidades establecidas, en este caso los valores a predecir se encuentran definidos en un conjunto de posibles valores como son : Infraestructura, Backend, FrontEnd y Base de Datos, teniendo en cuenta lo anterior, podemos evidenciar en la Figura 8 se implementa el algoritmo de clasificación con árboles de decisión en knime con los siguientes nodos:

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 30 de 52

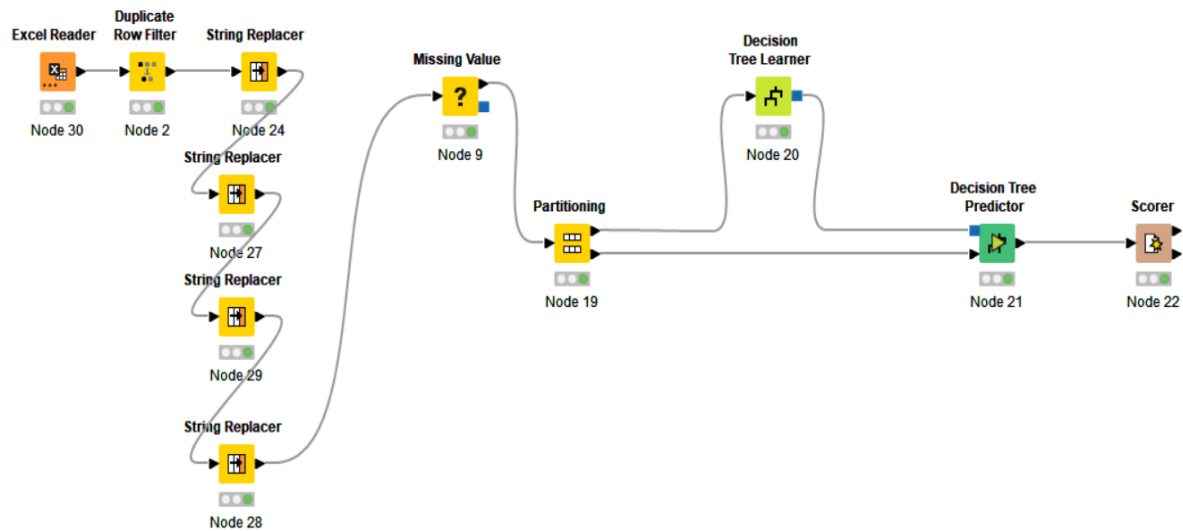



Figura 8. Modelo de Clasificación Arboles de decisión, Imagen Propia, 2023 Fuente. Knime Analytics Platform

Excel Reader: Cargamos el archivo que genero el modelo de agrupamiento resultado tomado del nodo desnormalizer en el cual se encuentran las habilidades técnicas de los candidatos junto con los 8 cluster asignados.

String Replacer: este nodo como su nombre lo indica reemplaza el valor de una cadena de string por otra cadena de string dada, en este caso se implementan 4 nodos con el fin de agrupar los 8 clusteres en los 4 grupos de habilidades según el resultado generado en el análisis de datos.


Partitioning: a partir de una tabla de entrada este nodo genera la división en dos particiones en este caso uno para el entrenamiento y otro para las pruebas de predicción, el valor del particionamiento es configurable en este caso 70 entrenamiento – 30 predicción

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 31 de 52

Decision Tree Learner: nodo utilizado para construir modelos de árboles de decisión a partir de conjuntos de datos de entrenamiento. El árbol de decisión es un modelo de aprendizaje automático que utiliza una estructura en forma de árbol para tomar decisiones o clasificar instancias en función de características o atributos. El resultado es un modelo de árbol de decisión que se puede utilizar para realizar predicciones en nuevos conjuntos de datos o evaluar la importancia de las variables predictoras.

Decision Tree Predictor: es utilizado para aplicar un modelo de árbol de decisión previamente construido a nuevos conjuntos de datos y realizar predicciones. Este nodo toma como entrada un modelo de árbol de decisión y un conjunto de datos de prueba, y produce como salida las predicciones basadas en el modelo. Se configura especificando las columnas objetivo y el modelo de árbol de decisión previamente construido. En este caso para la entrada se utiliza el árbol de decisión generado en el nodo anterior y los datos de entrada particionados en el 30% sin etiqueta, el resultado son los datos entrantes con la predicción de la etiqueta del grupo en el que quedaría cada candidato.

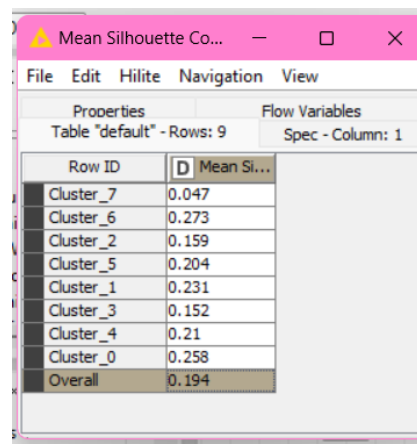
Scorer: se utiliza para evaluar el rendimiento de un modelo de aprendizaje automático utilizando métricas de evaluación. Proporciona una forma conveniente de calcular y visualizar métricas comunes, toma como entrada un conjunto de datos con las predicciones del modelo y las etiquetas reales (conocidas como el conjunto de datos de referencia). Luego, calcula las métricas de evaluación utilizando estas dos fuentes de datos y genera un informe detallado.

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 32 de 52

Semana 11: Métricas de los algoritmos implementados


Tarea 1.6: Generar métricas del algoritmo de agrupamiento.

En este punto es importante detallar con que precisión se agrupan y clasifican los datos generados, lo anterior se debe plasmar en métricas existentes implementadas en cada modelo, de acuerdo a lo anterior como se visualiza en la Figura 6 para el algoritmo de agrupamiento se usa el nodo Silhouette Coefficient como se mencionó en la generación del modelo de agrupamiento se utiliza este nodo que calcula el coeficiente de silueta para el resultado de agrupación siendo útil para evaluar el rendimiento, este funciona generando un cálculo por fila usando $(b-a)/\max(a,b)$ donde a es la distancia media dentro del grupo y b es la distancia media entre grupos al grupo más próximo, para el resultado general se calcula una tabla con la media de todos los coeficientes de silueta individuales, el rango de puntuación esta entre -1,0 y 1,0 entre mayor sea la puntuación más preciso es el agrupamiento generado.



Row ID	Mean Si...
Cluster_7	0.047
Cluster_6	0.273
Cluster_2	0.159
Cluster_5	0.204
Cluster_1	0.231
Cluster_3	0.152
Cluster_4	0.21
Cluster_0	0.258
Overall	0.194

Figura 9. Resultado Nodo Coeficiente de silueta, Imagen Propia, 2023 Fuente. Knime Analytics

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 33 de 52


Tarea 1.7: Generar métricas del algoritmo de clasificación.

En el cumplimiento de la tarea de generación de métricas con el fin de garantizar una buena clasificación del candidato con las habilidades técnicas en los grupos o áreas específicas como se evidencia en la Figura 7, se implementa el nodo Scorer que se encarga de comparar dos columnas por sus pares de valores de atributos mostrando la matriz de confusión es decir cuantas filas de cada atributo y su clasificación coinciden, sus configuraciones permiten seleccionar dos columnas para comparar en la primera columna seleccionada están los valores que se representan en las filas de la matriz de confusión y los valores de la segunda columna por las columnas de la matriz de confusión, esta contiene los que fueron correctamente clasificados, los verdaderos positivos, falsos positivos, verdaderos negativos, falsos negativos, precisión, recuperación, sensibilidad, especificidad, medida F, así como la precisión general y kappa de Cohen.

Cluster \ P...	Cluster_6	Cluster_2	Cluster_3	Cluster_7
Cluster_6	11	0	0	1
Cluster_2	0	16	0	0
Cluster_3	3	2	14	0
Cluster_7	0	0	1	12

Correct classified: 53	Wrong classified: 7
Accuracy: 88,333%	Error: 11,667%
Cohen's kappa (κ): 0,844%	

Figura 10. Resultado Nodo Scorer, Imagen Propia, 2023 Fuente. Knime Analytics Platform

<p>Información Privada</p> 	<p>PROCESO: INVESTIGACIÓN</p>	<p>IC-03 1922019</p>
	<p>SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA</p>	<p>Versión :2</p>
	<p>FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN</p>	<p>Página: 34 de 52</p>

Semana 12 -13- 14: Generación Smart Contract en Blockchain

Tarea 9.1 Implementación del Smart Contract

Implementamos el proyecto en Visual Studio Code instalando las siguientes extensiones:

Remix: Prueba y aseguramiento del Contrato Inteligente.

Solidity: Lenguaje para desarrollar en Contrato inteligente.

Python: Lenguaje para desarrollar el Backend.

Comenzamos con la instalación de Brownie por consola con pipx (instala Brownie en un entorno virtual y lo hace disponible directamente desde la línea de comandos.). ejecutando los siguientes comandos:

- Instalación pipx



```
python3 -m pip install --user pipx
```

```
python3 -m pipx ensurepath
```
- Instalación brownie usando pipx


```
pipx install eth-brownie
```
- Confirmación de la instalación de Brownie


```
$ brownie
```


Brownie - Python development framework for Ethereum.
- Instalación de Ganache a través de la página oficial de Truffle, Confirmamos la instalación a través del comando en cmd.

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 35 de 52

\$Ganache –versión

- Desarrollamos "deploy.py" es un archivo de script en Python que se utiliza para desplegar contratos inteligentes en una red blockchain utilizando la biblioteca de desarrollo de contratos inteligentes de Brownie. El script contiene la lógica necesaria para interactuar con la red blockchain, configurar las cuentas, establecer las estrategias de gas y finalmente desplegar el contrato inteligente en la red.
- Generamos la Programación del contrato inteligente con las siguientes funciones:
 - Crear Empleado: Crea el empleado como una transacción en el bloque de la red de blockchain
 - Crear Contrato: Le crea un contrato al empleado establecido
 - Obtener Empleado: Trae y/o consulta la información del empleado definido
 - Obtener Contrato: Trae y/o consulta la información del contrato laboral definido

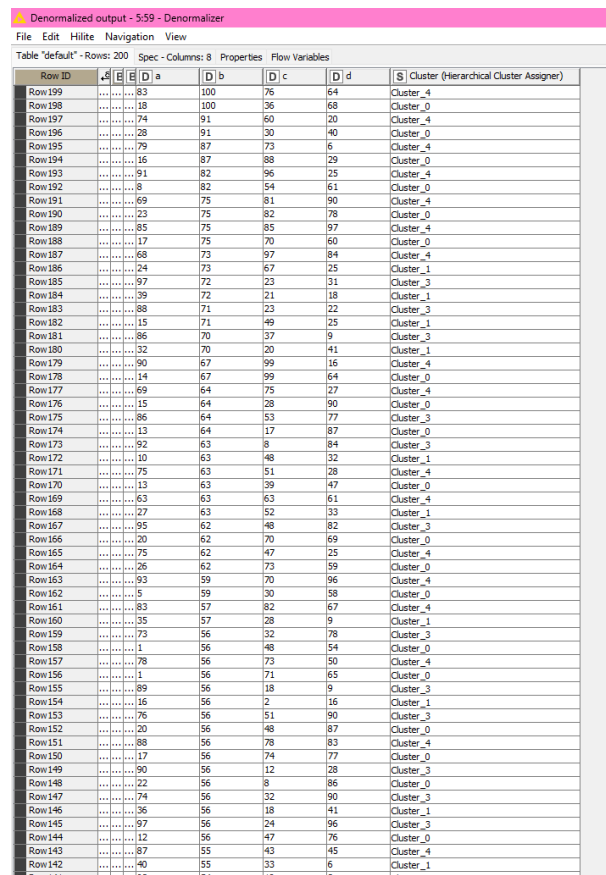
Detallamos la importancia de la creación del empleado y el contrato laboral ya que se realiza de forma automática cargando la información en un archivo plano.

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 36 de 52

8. RESULTADOS


- Resultados Algoritmo de Agrupamiento

Para el algoritmo de agrupamiento se cargan los datos generados en la semana 2 y 3, se ejecutan todos los nodos en orden teniendo en cuenta que todos queden ejecutados sin errores el resultado como se evidencia en la Figura 11 es el Excel con los datos de las habilidades técnicas y los clústeres asignados.



Row ID	a	b	c	d	Cluster (hierarchical Cluster Assigner)
Row199	83	100	76	64	Cluster_4
Row198	18	100	36	68	Cluster_0
Row197	74	91	60	20	Cluster_4
Row196	28	91	30	40	Cluster_0
Row195	79	87	73	6	Cluster_4
Row194	16	87	88	29	Cluster_0
Row193	91	82	96	25	Cluster_4
Row192	8	82	54	61	Cluster_0
Row191	69	75	81	90	Cluster_4
Row190	23	75	82	78	Cluster_0
Row189	85	75	85	97	Cluster_4
Row188	17	75	70	60	Cluster_0
Row187	88	73	97	84	Cluster_4
Row186	24	73	67	25	Cluster_1
Row185	97	72	23	31	Cluster_3
Row184	39	72	21	18	Cluster_1
Row183	88	71	23	22	Cluster_3
Row182	15	71	49	25	Cluster_1
Row181	86	70	37	9	Cluster_3
Row180	32	70	20	41	Cluster_1
Row179	90	67	99	16	Cluster_4
Row178	14	67	99	64	Cluster_0
Row177	69	64	75	27	Cluster_4
Row176	15	64	28	90	Cluster_0
Row175	86	64	53	77	Cluster_3
Row174	13	64	17	87	Cluster_0
Row173	92	63	8	84	Cluster_3
Row172	10	63	48	32	Cluster_1
Row171	75	63	51	28	Cluster_4
Row170	13	63	39	47	Cluster_0
Row169	63	63	63	61	Cluster_4
Row168	27	63	52	33	Cluster_1
Row167	95	62	48	82	Cluster_3
Row166	20	62	70	69	Cluster_0
Row165	75	62	47	25	Cluster_4
Row164	26	62	73	59	Cluster_0
Row163	93	59	70	96	Cluster_4
Row162	5	59	30	58	Cluster_0
Row161	83	57	82	67	Cluster_4
Row160	35	57	28	9	Cluster_1
Row159	73	56	32	78	Cluster_3
Row158	1	56	48	54	Cluster_0
Row157	78	56	73	50	Cluster_4
Row156	1	56	71	65	Cluster_0
Row155	89	56	18	9	Cluster_3
Row154	16	56	2	16	Cluster_1
Row153	76	56	51	90	Cluster_3
Row152	20	56	48	87	Cluster_0
Row151	88	56	78	83	Cluster_4
Row150	17	56	74	77	Cluster_0
Row149	90	56	12	28	Cluster_3
Row148	22	56	8	86	Cluster_0
Row147	74	56	32	90	Cluster_3
Row146	36	56	18	41	Cluster_1
Row145	97	56	24	96	Cluster_3
Row144	12	56	47	76	Cluster_0
Row143	87	55	43	45	Cluster_4
Row142	40	55	33	6	Cluster_1
Row141	83	64	18	2	Cluster_3

Figura 11. Resultado Algoritmo de Agrupamiento, Imagen Propia, 2023 Fuente. Knime

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 37 de 52

- Resultados Análisis de Datos

Posterior a esto se procede con el análisis de datos del resultado del paso anterior en el cual se evidencian los 8 grupos, se genera un análisis respectivo mencionado en la semana 7 y 8 realizando la agrupación en 4 grupos enfocados en las habilidades de la siguiente manera:

Cluster 2 y Cluster 0: a: Base de Datos


Cluster 6 y Cluster 4: b: Infraestructura

Cluster 1 y Cluster 3: c: Backend

Cluster 5 y Cluster 7: d : FrontEnd

- Resultados Algoritmo de Clasificación

Una vez se obtienes los grupos enmarcados de las habilidades técnicas se procede a cargar el Excel generado del modelo de agrupamiento al modelo de clasificación con el fin de que el modelo entrene con los datos establecido y genere la predicción de forma correcta, se ejecutan todos los nodos verificando que se ejecuten exitosamente, al finalizar la ejecución que no tarda mas de 1 segundo como se visualiza en la Figura 12 se genera la tabla de decisión del nodo tree predictor que predice el 30% de los datos inicialmente cargados.

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 38 de 52


Classified Data - 4:21 - Decision Tree Predictor

File Edit Hilitte Navigation View

Table "default" - Rows: 60 Spec - Columns: 6 Properties Flow Variables


Row ID	D a	I b	D c	I d	S Cluster	S Predict...
Row0	83	100	76	64	Cluster_6	Cluster_6
Row1	18	100	36	68	Cluster_2	Cluster_2
Row2	74	91	60	20	Cluster_6	Cluster_6
Row3	28	91	30	40	Cluster_2	Cluster_2
Row6	91	82	96	25	Cluster_6	Cluster_6
Row7	8	82	54	61	Cluster_2	Cluster_2
Row11	17	75	70	60	Cluster_2	Cluster_2
Row16	88	71	23	22	Cluster_3	Cluster_3
Row17	15	71	49	25	Cluster_3	Cluster_3
Row18	86	70	37	9	Cluster_3	Cluster_6
Row19	32	70	20	41	Cluster_3	Cluster_2
Row20	90	67	99	16	Cluster_6	Cluster_6
Row21	14	67	99	64	Cluster_2	Cluster_2
Row22	69	64	75	27	Cluster_6	Cluster_6
Row24	86	64	53	77	Cluster_3	Cluster_6
Row31	27	63	52	33	Cluster_3	Cluster_3
Row37	5	59	30	58	Cluster_2	Cluster_2
Row38	83	57	82	67	Cluster_6	Cluster_6
Row40	73	56	32	78	Cluster_3	Cluster_3
Row45	16	56	2	16	Cluster_3	Cluster_3
Row46	76	56	51	90	Cluster_3	Cluster_6
Row47	20	56	48	87	Cluster_2	Cluster_2
Row48	88	56	78	83	Cluster_6	Cluster_6
Row50	90	56	12	28	Cluster_3	Cluster_3
Row52	74	56	32	90	Cluster_3	Cluster_3
Row53	36	56	18	41	Cluster_3	Cluster_2
Row54	97	56	24	96	Cluster_3	Cluster_3
Row57	40	55	33	6	Cluster_3	Cluster_3
Row58	93	54	18	3	Cluster_3	Cluster_3
Row60	72	54	9	20	Cluster_3	Cluster_3
Row65	5	53	24	90	Cluster_2	Cluster_2
Row66	71	52	33	6	Cluster_3	Cluster_3
Row67	34	52	89	32	Cluster_7	Cluster_3
Row75	29	51	19	3	Cluster_3	Cluster_3
Row83	43	47	70	83	Cluster_2	Cluster_2
Row89	52	45	4	83	Cluster_2	Cluster_2
Row90	48	45	62	25	Cluster_7	Cluster_7
Row93	50	45	54	51	Cluster_2	Cluster_2
Row95	56	45	92	74	Cluster_2	Cluster_2
Row96	55	45	99	22	Cluster_7	Cluster_7
Row107	49	43	83	47	Cluster_7	Cluster_7
Row111	60	42	63	17	Cluster_7	Cluster_7
Row115	57	39	95	20	Cluster_7	Cluster_7
Row121	52	39	40	63	Cluster_2	Cluster_2
Row123	53	39	4	55	Cluster_2	Cluster_2
Row129	55	35	94	34	Cluster_7	Cluster_7
Row132	48	35	34	51	Cluster_2	Cluster_2
Row141	51	33	42	33	Cluster_3	Cluster_3
Row147	54	31	19	68	Cluster_2	Cluster_2
Row156	61	28	67	45	Cluster_6	Cluster_7
Row161	26	27	87	28	Cluster_7	Cluster_7
Row165	14	24	25	62	Cluster_7	Cluster_7
Row166	92	24	93	40	Cluster_6	Cluster_6
Row172	61	20	21	19	Cluster_7	Cluster_7
Row178	73	17	46	37	Cluster_6	Cluster_6
Row180	98	16	100	24	Cluster_6	Cluster_6
Row185	13	14	99	34	Cluster_7	Cluster_7
Row197	6	11	53	23	Cluster_7	Cluster_7

Figura 12. Resultado Nodo Tree Predictor, Imagen Propia, 2023 Fuente. Knime

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 39 de 52

De igual manera de acuerdo con la Semana 11 Tarea 1,7 se evidencia la matriz de confusión en la cual muestra que el modelo predice con un 83,3% de efectividad los datos cargados, por lo cual se toma como un modelo confiable y preciso para la clasificación de los candidatos en las áreas de interés, adicionalmente de acuerdo con la Figura 13, se generan pruebas del árbol generado en el árbol de decisión visualizado en la Figura 14.

Las pruebas realizadas se generaron con la carga de 4 candidatos sin etiqueta en el cual se evidenciaba que el algoritmo de clasificación lo catalogaba en algún grupo designado siguiendo la ruta del árbol de decisión se evidencia que de las 4 cargas el modelo clasifica con una precisión del 100% dos candidatos.

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 40 de 52

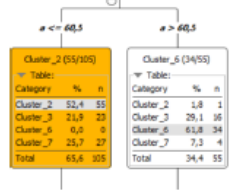
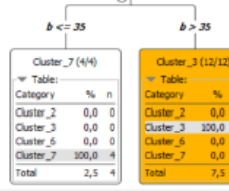
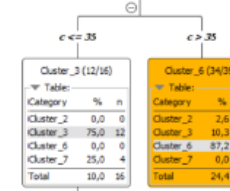
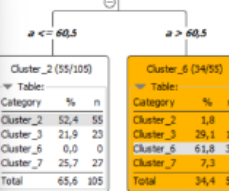
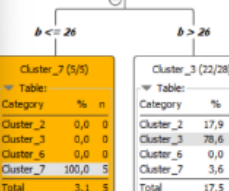
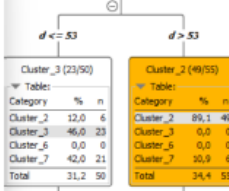
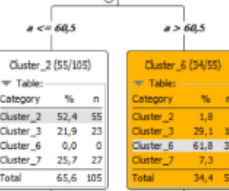
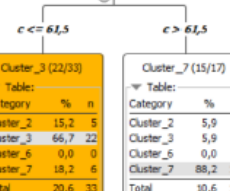
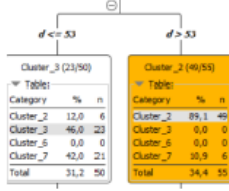
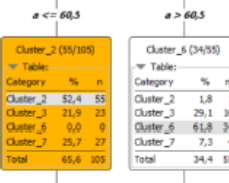
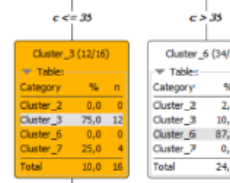
Skill 1	Skill 2	Skill 3	Skill 4
85	14	3	66
Cluster 6	Cluster 7	Cluster 3	No muestra
			
34	77	55	5
Cluster 2	Cluster 3	No muestra	Cluster 3
			
25	12	63	31
Cluster 2	No muestra	Cluster 7	Cluster 3
			
78	15	47	96
Cluster 6	Cluster 2	Cluster 6	No muestra
			

Figura 13. Pruebas Modelo Clasificación, Imagen Propia, 2023

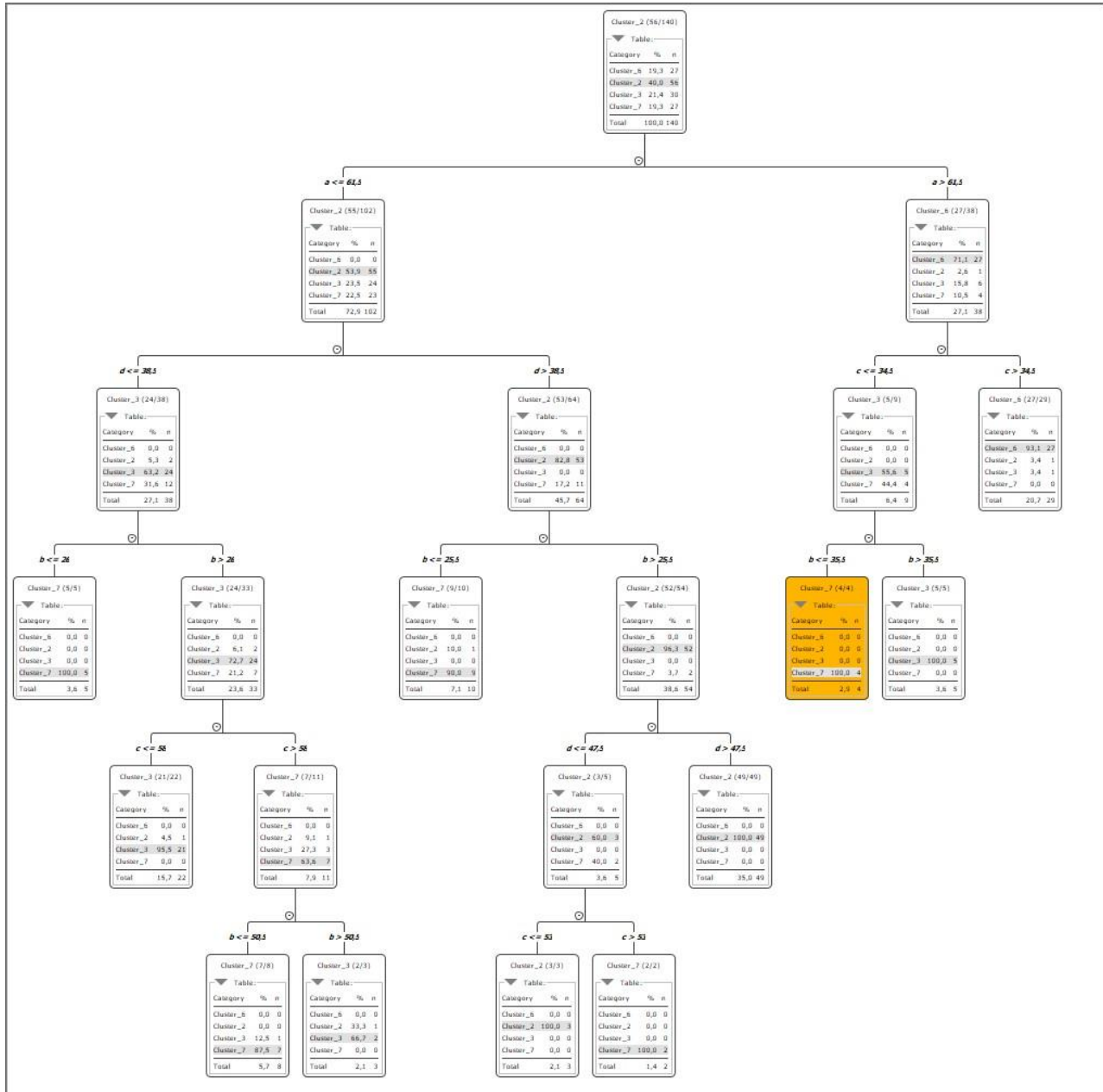

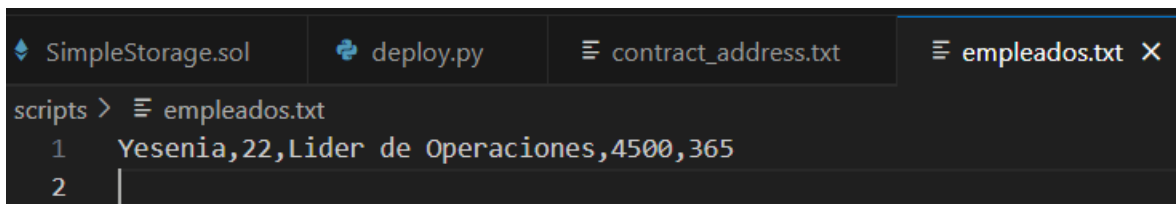


Figura 14. Resultado Árbol de Decisión, Imagen Propia, 2023 Fuente. Knime Analytics Platform

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 42 de 52

- Resultados Smart Contract Blockchain

En la generación del contrato laboral para un candidato, Como vemos en la Figura 15 se carga el archivo empleado.txt en el cual contiene la información del candidato, así como la duración en meses y el salario establecido.



```

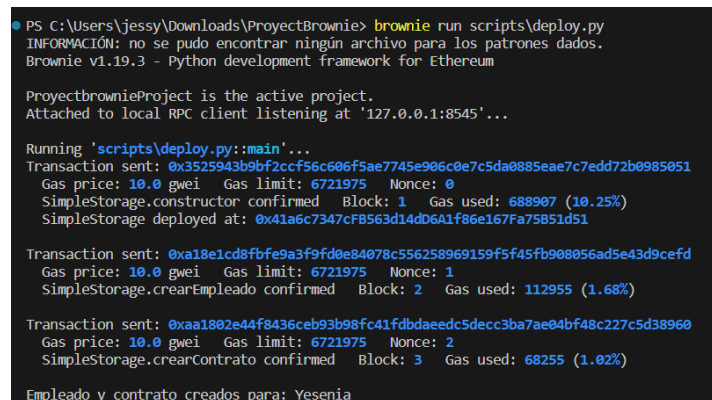
SimpleStorage.sol  deploy.py  contract_address.txt  empleados.txt X
scripts > empleados.txt
1  Yesenia,22,Lider de Operaciones,4500,365
2

```

Figura 15. Archivo plano Información Empleado, Imagen Propia, 2023 Fuente. Visual Studio

Code

Una vez cargado el archivo se despliega el archivo deploy.py en el cual como se visualiza en la Figura 16, se genera el contrato inteligente con dos transacciones que indican la creación del empleado y del contrato laboral para el empleado cargado.



```

PS C:\Users\jessy\Downloads\ProjectBrownie> brownie run scripts\deploy.py
INFORMACIÓN: no se pudo encontrar ningún archivo para los patrones dados.
Brownie v1.19.3 - Python development framework for Ethereum

ProjectbrownieProject is the active project.
Attached to local RPC client listening at '127.0.0.1:8545'...

Running 'scripts\deploy.py::main'...
Transaction sent: 0x3525943b9bf2ccf56c606f5ae7745e906c0e7c5da0885eae7c7edd72b0985951
Gas price: 10.0 gwei Gas limit: 6721975 Nonce: 0
SimpleStorage.constructor confirmed Block: 1 Gas used: 688907 (10.25%)
SimpleStorage deployed at: 0x41a6c7347cFB563d14dD6A1f86e167Fa75851d51

Transaction sent: 0xa18e1cd8fbfe9a3f9fd0e84078c556258969159f5f45fb908056ad5e43d9cefd
Gas price: 10.0 gwei Gas limit: 6721975 Nonce: 1
SimpleStorage.crearEmpleado confirmed Block: 2 Gas used: 112955 (1.68%)


Transaction sent: 0xaa1802e44f8436ceb93b98fc41fdbdae5decc3ba7ae04bf48c227c5d38960
Gas price: 10.0 gwei Gas limit: 6721975 Nonce: 2
SimpleStorage.crearContrato confirmed Block: 3 Gas used: 68255 (1.02%)

Empleado y contrato creados para: Yesenia

```

Figura 16. Resultado Creación Smart Contract, Imagen Propia, 2023 Fuente. Visual Studio

Code

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 43 de 52

Esta información debe evidenciarse consignada en la blockchain local Ganache en la Figura 17 podemos ver el contrato desplegado con 3 transacciones vistas en la Figura 17

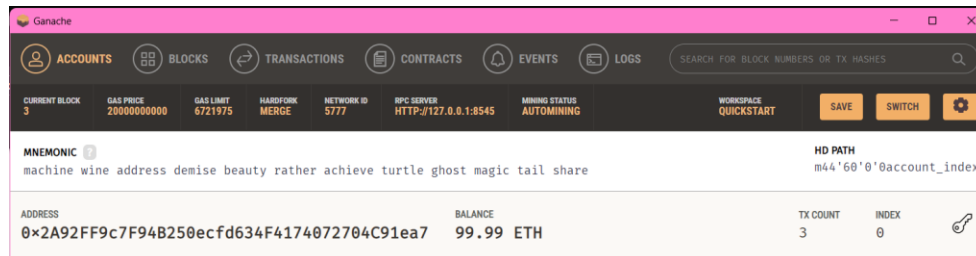


Figura 17. Resultado Blockchain Local, Imagen Propia, 2023 Fuente. Ganache

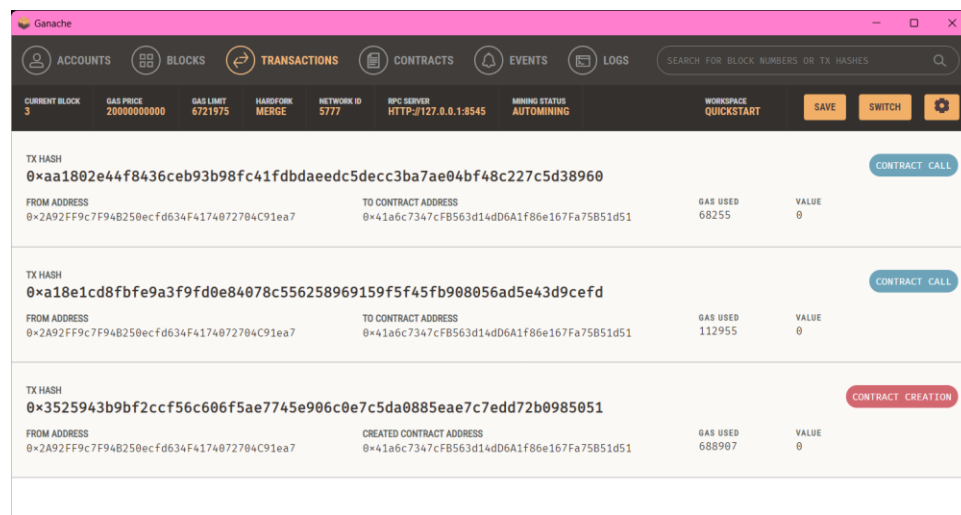




Figura 18. Transacciones Blockchain Local, Imagen Propia, 2023 Fuente. Ganache

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 44 de 52


9. CONCLUSIONES Y RECOMENDACIONES

El proceso de diseño de un algoritmo de agrupamiento se llevó a cabo, mediante la implementación de un agrupamiento jerárquico en Knime y la incorporación de un conjunto de datos aleatorios representados por las habilidades técnicas que posee un candidato que se postula a una vacante y el nivel de conocimiento que presenta por cada una de ellas, donde se evidencia el agrupamiento con 8 clústeres a partir de 200 registros, de esta manera se evidenció la categorización de los posibles candidatos enfocados en su nivel de conocimiento con habilidades técnicas específicas, adicional los nodos implementados en Knime facilitaron la visualización y comprensión de los resultados logrando ser interpretados e identificando las habilidades técnicas que más se repetían en cada cluster, a través de un análisis desde Excel. Este proceso tiene el potencial de informar estrategias de agrupamiento de habilidades y asignación de roles de manera más precisa en entornos laborales o académicos.

A partir del diseño anterior se implementó una nueva cadena de bloques en Knime donde se diseñó un algoritmo de clasificación, haciendo uso del algoritmo de árbol de decisión que permitió realizar una clasificación de un candidato según su nivel en cada habilidad técnicas, de manera que se insertaron los datos agrupados del flujo con el algoritmo de agrupamiento, este proceso permitió conocer la capacidad del modelo para clasificar correctamente de acuerdo con los datos proporcionados, proporcionando una base sólida para la toma de decisiones. Los candidatos fueron consistentemente asignados a clusters específicos según su nivel más alto en una habilidad, lo que sugiere una clasificación coherente y confiable.

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 45 de 52

La construcción del contrato inteligente como la etapa final del proceso ofrece transparencia y eficiencia en los procesos de selección de las empresas, de manera se mantiene la trazabilidad de las validaciones necesarias para que una persona sea contratada, pero ahorrando tiempos y recursos, adicional se evidencia que es un proceso seguro ya que el blockchain utiliza criptografía para garantizar la seguridad de los datos por lo que solo se comparte la información necesaria. Esta etapa representa la evolución en la gestión de recursos humanos, optimizando la toma de decisiones y el aumento de confianza en los procesos de selección.

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 46 de 52

10. REFERENCIAS BIBLIOGRÁFICAS

Aydoğan, R., Marsa-Maestre, I., Klein, M., & M. Jonker , C. (2018). A Machine Learning Approach for Mechanism Selection in Complex Negotiations. *Journal of Systems Science and Systems Engineering*, 134-155. doi:<https://doi.org/10.1007/s11518-018-5369-5>


Barrientos-Martinez, R., Cruz-Ramirez, N., Acosta-Mesa, H.-G., Rabatte-Suárez, I., Gogeoascoechea-Trejo, M.-d.-C., Pavón-León, P., & Blázquez-Morales, S.-L. (2009). Árboles de decisión como herramienta en el diagnóstico médico. *Revista médica de la Universidad Veracruzana*, 19.24.

Bartoletti, M., & Pompianu, L. (2017). *An empirical analysis of smart contracts*. Università degli Studi di Cagliari, Cagliari, Italy. doi:<https://arxiv.org/abs/1703.06322>

Beltran-Pardo, M., & Sevillano-Jaén, F. (2013). *Cloud Computing, tecnología y negocio*. Ediciones Paraninfo, S.A.

Bojorque-Chasi, R.-X. (2020). *Clustering de sistemas de recomendación mediante técnicas de factorización matricial*. Universidad Politécnica de Madrid. Madrid: E.T.S.I. de Sistemas Informáticos (UPM). doi:<https://oa.upm.es/58156/>

Caballero, J., Martinez, C., & Ballesteros, F. (n.d.). *INSPECCIONES POR ULTRASONIDOS ORIENTADAS POR LOS DATOS PARA MATERIALES COMPUESTOS*. Madrid: REVISTA DE LA SOCIEDAD ESPAÑOLA DE MATERIALES.

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 47 de 52

Chiavenato, A. (2011). *Administración de recursos humanos: El capital humano de las organizaciones*. University of Los Angeles: McGRAW-HILL/INTERAMERICANA.

Chiavenato, I. (2009). *Gestión de Recursos Humanos*. (E. E. Ltda, Ed.) The McGraw-Hill .

Corral, F. (2007). *Reclutamiento y Selección por Competencias*. EScuela de Negocios. Caracas-Venezuela: eoiamérica.

Diaz-Romero, S., & Sanyer Mosquera, W. (2021). “*Selección de candidatos para encuestas mediante técnicas de machine learning*”. Guayaquil - Ecuador:

<https://www.dspace.espol.edu.ec/handle/123456789/54068>.

doi:<https://www.dspace.espol.edu.ec/handle/123456789/54068>

Emilio. (2020, Julio 28). *Introducción a Knime*. Retrieved from TodoBI Business Intelligence, stratebi open business intelligence: <https://todobi.com/introduccion-a-kine/>


García-Bermejo, P. (2022). *Data Science y KNIME, combinación perfecta para el éxito en la toma de decisiones*. Universitat Oberta de Catalunya (UOC).

doi:<http://hdl.handle.net/10609/138786>

García-Cambronero, C., & Gomez-Moreno, I. (2006). *Algoritmos de aprendizaje: knn & kmeans*. Universidad Carlos III de Madrid, Madrid.

Golosova, J., & Romanovs, A. (2018, December 30). *IEEE Xplore*.

doi:10.1109/AIEEE.2018.8592253

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 48 de 52

Gonzalez-Perez, M.-A. (2018, Noviembre 17). Blockchain: más que transparencia y descentralización. doi:<http://hdl.handle.net/10784/24560>


Guerrero Barrera, E. (2017). *"PRINCIPALES ESTRATEGIAS DE AFRONTAMIENTO DE ESTRÉS LABORAL EN LOS PROCESOS DE RECLUTAMIENTO MASIVO, SEGÚN LA PERCEPCIÓN DE UN GRUPO DE ENCARGADOS DE DOTACIÓN DE PERSONAL."*. GUATEMALA DE LA ASUNCIÓN: UNIVERSIDAD RAFAEL LANDÍVAR.

Hernández-Martínez, L. (2020). *El reclutamiento 4.0 en la era digital: La selección de talento y el marketing digital*. Universidad Zaragoza, Departamento de Informática e Ingeniería de Sistemas, Área de Lenguajes y Sistemas Informáticos. ZAGUAN Repositorio Institucional de Documentos. doi:<https://zaguan.unizar.es/record/101929>

Hinestroza-Ramírez, D. (2018). *El Machine Learning a través de los tiempos, y los aportes a la humanidad*. Universidad Libre Seccional de Pereira, Pereira. doi:<https://hdl.handle.net/10901/17289>

IEEE. (2016). *IEEE Xplore*. doi:10.1109/SP.2016.55.

Jacanamejoy-Jamioy, C.-A., Forero-Vargas, M.-G., & Tavera-Bucurú, W. (2019). *NUEVO MÉTODO PARA EL SEGUIMIENTO DE LA EVOLUCIÓN DE LOS ESTUDIANTES BASADOS EN TÉCNICAS DE AGRUPAMIENTO*. Universidad de Ibagué. Ibagué, Colombia: eiei Encuentro Internacional de Educación en Ingeniería ACOFI. doi:<https://doi.org/10.26507/ponencia.37>

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 49 de 52

Jumbo-Quichimbo, G.-N. (2019). *La inteligencia artificial y su impacto en los procesos de selección*. Universidad Internacional de la Rioja Máster Universitario en Dirección y Gestión de Recursos Humanos. Loja-Ecuador: Reunir, Repositorio Digital. doi:<https://reunir.unir.net/handle/123456789/8166>


Knirsch, F., & Unterweger, A. (2019, March 11). Implementing a blockchain from scratch: why, how, and what we learned. doi:<https://link.springer.com/article/10.1186/s13635-019-0085-3>

LegalTech news. (2018). A Brief History of BlockcHain. *ALM*, 1 - 2. doi:<chrome-extension://efaidnbmninnibpcajpegclclefindmkaj/https://pbwt2.gjassets.com/content/uploads/2018/05/010051804-Patterson2.pdf>

Lopez Gumucio, J. (2010). LA SELECCIÓN DE PERSONAL BASADA EN COMPETENCIAS Y SU RELACIÓN CON LA EFICACIA ORGANIZACIONAL. *Sistema de Información Científica Redalyc*. Retrieved from <https://www.redalyc.org/articulo.oa?id=425941230007>

Lotito Catino, F. (2015). Test psicológicos y entrevistas: usos y aplicaciones claves en el proceso de selección e integración de personas a las empresas. *RAN Revista Academia & Negocios*. Retrieved from <https://revistas.udec.cl/index.php/ran/article/view/3005/3101>

Maisueche-Cuadrado, A. (2019). *Utilización del Machine Learning en la industria 4.0*. Universidad de Valladolid. Escuela de Ingenierías Industriales. doi:<http://uvadoc.uva.es/handle/10324/37908>

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 50 de 52

Monje López, R. (2021). *Algoritmo Jerárquico Aglomerativo Paralelo para la*. UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS. CUERNAVACA: UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS. Retrieved from <http://riaa.uaem.mx/handle/20.500.12055/1672>


Pascual, D., Pla, F., & Sánchez, S. (2007). *Algoritmos de agrupamiento. Método Informáticos Avanzados*. Universidad de Oriente, Santiago de Cuba.

Pastor-Martínez, M.-Á. (2020). *RECLUTAMIENTO 4.0*. Universidad de Jaén, Facultad de Ciencias Sociales y Jurídicas. doi:<https://hdl.handle.net/10953.1/12816>

Puyol-Moreno, J. (2014). *UNA APROXIMACIÓN A BIG DATA - AN APPROACH TO BIG DATA*. *Revista de deRecho UNed*.

Rebafka, T. (2023). *Model-based clustering of multiple networks with a hierarchical algorithm*. Sorbonne University. Research Square. doi:<https://doi.org/10.21203/rs.3.rs-2494480/v1>

Rojko, A. (2017). *Industry 4.0 Concept: Background and Overview*. *International journal of interactive mobile technologies*. Retrieved from <https://web.s.ebscohost.com/abstract?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=18657923&AN=124305798&h=36NYn1k0amLuIFp6mHgUkdEUF8d26sm9DweV8gmEdfqGrVVzIcdwrQcobYNDvfGTvZYZ%2fcDvHCQeD%2fvsaZkw2w%3d%3d&crl=c&resultNs=AdminWebAuth&resultLoca>

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 51 de 52

Romero, M. (2023). *PROCESO DE SELECCIÓN DE PERSONAL: IMPLICANCIAS DE LA INCORPORACIÓN DE TECNOLOGÍA EN EL PROCESO*. Montevideo. doi:<https://hdl.handle.net/20.500.12008/39061>


Salazar, J., & Silvestre, S. (2017). *Internet de las cosas*. České vysoké učení technické v Praze Fakulta elektrotechnická.

Salgado, J., & Peiro, J. (2008). Psicología del trabajo, las organizaciones y los recursos humanos en España. *Papeles del Psicólogo*. Retrieved from <http://www.redalyc.org/articulo.oa?id=77829101>

Sandoval, L.-J. (2018). ALGORITMOS DE APRENDIZAJE AUTOMÁTICO PARA ANÁLISIS Y PREDICCIÓN DE DATOS. *REVISTA TECNOLÓGICA N° 11*. doi:<http://redicces.org.sv/jspui/handle/10972/3626>

Sapriani, G., Iwan, P., Windo, & Erlinawaty. (2021). Student Grouping Based on Final Exam Values of the Courses with the K-Means Classification Method Using KNIME. *Internet of Things and Artificial Intelligence Journal*, 114-119. doi:<https://doi.org/10.31763/iota.v1i2.378>

Su, J., & Zhang, H. (2006). *A Fast Decision Tree Learning Algorithm*. University of New Brunswick, NB, Canada, E3B 5A3.

<p>Información Privada</p> 	PROCESO: INVESTIGACIÓN	IC-03 1922019
	SUBPROCESO: GESTIÓN DE LOS PROGRAMAS DE INICIACIÓN CIENTÍFICA	Versión :2
	FORMATO: TRABAJO DE GRADO - PROYECTOS DE INVESTIGACIÓN	Página: 52 de 52

Tasende, I. (2020). Blockchain y arbitraje: un nuevo enfoque en la resolución de disputas. *Revista de Derecho. Publicacion Arbitrada de la Universidad Catolica del Uruguay*, 2-3. doi:<https://doi.org/10.22235/rd.vi22.2127>

Valencia-Ramirez, J.-P. (2019). Contratos inteligentes. *Revista RITI Investigación en Tecnologías de la Información*, 1-10. doi:<https://doi.org/10.36825/RITI.07.14.001>

Zahra, S., Munstansar-Ali, G., Asra, K., Muhammad-Awais, A., Usman, N., & Prugel-Bennett, A. (2015). Novel centroid selection approaches for KMeans-clustering based recommender systems. *Information Sciences* Volume 320, 156-189. doi:<https://doi.org/10.1016/j.ins.2015.03.062>