



**UNIVERSIDAD TÉCNICA DE COTOPAXI**  
**FACULTAD DE CIENCIAS DE LA INGENIERÍA Y APLICADAS.**  
**CARRERA DE INGENIERÍA EN INFORMÁTICA Y SISTEMAS**  
**COMPUTACIONALES.**

**PROYECTO DE INVESTIGACIÓN**

**“MODELO BASADO EN MACHINE LEARNING PARA LA PREDICCIÓN DE  
VARIABLES EMPLEADAS EN EL SECADO DE SEMILLA EN HORNO  
MICROONDAS”**

Proyecto de Titulación presentado previo a la obtención del Título de Ingenieras en Informática y Sistemas Computacionales.

**Autoras:**

Cárdenas Quinapaxi Silvia Carolina.

Vela Pérez Vanessa Estefanía.

**Tutora:**

Ing. Bravo Mullo Silvia Jeaneth.

LATACUNGA-ECUADOR

AGOSTO-2019

## DECLARACIÓN DE AUTORÍA



Universidad  
Técnica de  
Cotopaxi



Ingeniería  
Informática Y Sistemas  
Computacionales

### DECLARACIÓN DE AUTORÍA

Nosotras, **CÁRDENAS QUINAPAXI SILVIA CAROLINA** y **VELA PÉREZ VANESSA ESTEFANÍA** declaramos ser autoras del presente proyecto de investigación: **“MODELO BASADO EN MACHINE LEARNING PARA LA PREDICCIÓN DE VARIABLES EMPLEADAS EN EL SECADO DE SEMILLAS EN HORNO MICROONDAS**, siendo la **ING. SILVIA JEANETH BRAVO MULLO** tutora del presente trabajo; y exímolo expresamente a la Universidad Técnica de Cotopaxi y a sus representantes legales de posibles reclamos o acciones legales.

Además, certificamos que las ideas, conceptos, procedimientos y resultados vertidos en el presente trabajo investigativo, son de nuestra exclusiva responsabilidad.

.....  
**Cárdenas Quinapaxi Silvia Carolina.**

C.C: 172317065-8.

.....  
**Vela Pérez Vanessa Estefanía.**

C.C: 180466449-6.

## APROBACIÓN DEL TRIBUNAL DE TITULACIÓN



Universidad  
Técnica de  
Cotopaxi



Ingeniería  
Informática Y Sistemas  
Computacionales

### APROBACIÓN DEL TRIBUNAL DE TITULACIÓN

En calidad de Tribunal de Lectores, aprueban el presente Informe de Investigación de acuerdo a las disposiciones reglamentarias emitidas por la Universidad Técnica de Cotopaxi, y por la FACULTAD de Ciencias de la Ingeniería y Aplicadas; por cuanto, las postulantes: Cárdenas Quinapaxi Silvia Carolina y Vela Pérez Vanessa Estefanía con el título del Proyecto de titulación: "Modelo basado en Machine Learning para la predicción de variables empleadas en el secado de semillas en horno microondas" han considerado las recomendaciones emitidas oportunamente y reúne los méritos suficientes para ser sometido al acto de Sustentación de Proyecto.

Por lo antes expuesto, se autoriza realizar los empastados correspondientes, según la normativa institucional.

Latacunga, 19 de Julio de 2019

Para constancia firman:

**Lector 1 (Presidente)**

**Nombre:** Ing. Karla Cantuña.  
**CC:** 050230511-3

**Lector 2**

**Nombre:** Ing. José Cadena  
**CC:** 050155279-8

**Lector 3**

**Nombre:** Ing. Oscar Guaypatin.  
**CC:** 180282943-3

## AVAL DEL TUTOR DE PROYECTO DE TITULACIÓN



Universidad  
Técnica de  
Cotopaxi



Ingeniería  
Informática Y Sistemas  
Computacionales

### AVAL DEL TUTOR DE PROYECTO DE TITULACIÓN

En calidad de Tutora del Trabajo de Investigación sobre el título:

**“MODELO BASADO EN MACHINE LEARNING PARA LA PREDICCIÓN DE VARIABLES EMPLEADAS EN EL SECADO DE SEMILLAS EN HORNO MICROONDAS”**, de Cárdenas Quinapaxi Silvia Carolina y Vela Pérez Vanessa Estefanía de la carrera de **Ingeniería en Informática y Sistemas Computacionales**, considero que dicho Informe Investigativo cumple con los requerimientos metodológicos y aportes científico-técnicos suficientes para ser sometidos a la evaluación del Tribunal de Validación de Proyecto que el Consejo Directivo de la Facultad de **Ciencias de la Ingeniería y Aplicadas** de la Universidad Técnica de Cotopaxi designe, para su correspondiente estudio y calificación.

Latacunga, Julio 2019.

.....  
Ing. Silvia Jeaneth Bravo Mullo.

C.C. 050243712-2

Tutora de Titulación.

## **AGRADECIMIENTO**

Agradezco a Dios por bendecirme en mi vida, por guiarme a lo largo de mi existencia, ser el apoyo y la fortaleza en aquellos momentos de dificultad y debilidad, a mis padres por su comprensión, paciencia y el apoyo incondicional y sobre todo por confiar en mí, ya que ellos fueron el pilar fundamental para cumplir mi meta dándome fuerzas para poder seguir adelante, a mi hermano por estar siempre presente acompañándome y por el apoyo moral que me brindó a lo largo de esta etapa y a toda mi familia que siempre me brindaron su apoyo y confianza en toda mi vida estudiantil.

A mis profesores que me brindaron conocimientos durante la carrera universitaria, de manera especial a la Ing. Silvia Bravo y al Phd. Ángel Hernández, a mi familia y amigos quienes me dieron su apoyo.

*Silvia Carolina Cárdenas Quinapaxi.*

## **DEDICATORIA**

A mis padres Ángel y Gladys quienes con su amor, paciencia y esfuerzo me han permitido llegar a cumplir hoy un sueño más, gracias por inculcar en mí el ejemplo de esfuerzo y valentía de no temer las adversidades porque Dios está conmigo siempre.

A mi hermano Ángel Fernando quien me motivó y me impulsó a seguir adelante y no me rinda, a toda mi familia por sus consejos, palabras de aliento y su cariño que hicieron de mí una mejor persona, cumpliendo mi sueño anhelado.

Finalmente quiero dedicar esta tesis a todas mis amigas, por apoyarme cuando más las necesito, por extender su mano en momentos difíciles y por el amor brindado cada día, de verdad mil gracias Mamita María, siempre te llevo en mi corazón.

*Silvia Carolina Cárdenas Quinapaxi.*

## **AGRADECIMIENTO**

Agradezco en primer lugar a Dios, por darme la fortaleza para cumplir esta gran meta, a mis padres que son mi pilar fundamental, mi inspiración, mi motivación para seguir adelante, que siempre estuvieron conmigo apoyándome en todo especialmente cuando más los necesitaba y a mi hermana que nunca dudo de mi y estuvo siempre a mi lado ayudándome en todo lo que estaba a su alcance.

Agradecer también a quien fue guía principal, mi tutora de titulación Ing. Silvia Bravo, que con su apoyo, su paciencia y su tiempo se pudo desarrollar satisfactoriamente este proyecto de investigación y a mis docentes quienes contribuyeron con sus conocimientos en toda mi etapa universitaria.

*Vanessa Estefanía Vela Pérez*

## **DEDICATORIA**

El presente trabajo, que representa dedicación y esfuerzo, está dedicado a mis padres y a mi hermana, quienes aparte de ser mi apoyo fundamental para seguir adelante, nunca dudaron de mí y siempre estuvieron conmigo durante toda mi etapa universitaria. Con sus palabras de aliento y sus sabios consejos, me enseñaron que nada es fácil pero tampoco imposible, que cuando se anhela algo hay que luchar hasta conseguirlo, es por eso que fueron quienes cada día me impulsaron para cumplir esta meta que nos propusimos los cuatro como familia y que hoy por hoy lo logramos a pesar de todas las adversidades que se han presentado durante el trayecto de esta etapa.

*Vanessa Estefanía Vela Pérez.*

## **UNIVERSIDAD TÉCNICA DE COTOPAXI**

### **FACULTAD DE CIENCIAS DE LA INGENIERÍA Y APLICADAS**

**TITULO:** “MODELO BASADO EN MACHINE LEARNING PARA LA PREDICCIÓN DE VARIABLES EMPLEADAS EN EL SECADO DE SEMILLA EN HORNO MICROONDAS”

**Autores:** Cárdenas Quinapaxi Silvia Carolina.

**Vela Pérez Vanessa Estefanía.**

#### **RESUMEN**

El secado de semillas implica la reducción del contenido de humedad a niveles recomendados, utilizando técnicas que no deterioren su viabilidad, eviten el deterioro, calentamiento e infestación durante su almacenamiento. El presente proyecto, se basa en desarrollar un modelo basado en Machine Learning para la predicción de las variables utilizadas en el secado de semillas, ya que no existe un modelo relacionado a este tipo de investigación, por ende no se puede obtener un buen rendimiento en el secado, lo cual se invierte recursos de tiempo, económicos y materiales.

Para el desarrollo del proceso de predicción se utilizaron técnicas predictivas como son: Redes Neuronales y Árboles de decisión, además, se empleó una metodología que involucra el método experimental y de prototipo. Adicionalmente, se utilizó el lenguaje de programación Python con la herramienta Skylearn, esta última herramienta permitió emplear la red neuronal MLPRegressor, la misma que fue calibrada con la siguiente característica principal, es decir se empleó 5 capas de entrada y 5 capas de salida, de igual manera para la predicción con árboles de decisión se empleó RandomForestRegressor, la misma que se calibró bajo la característica principal que en este caso fue el empleo de 10 estimadores y un estado aleatorio de 42. MLPRegressor alcanzó las siguientes predicciones, para el tiempo de secado: 99.18%, consumo de energía: 98.17% y tasa de germinación: 78.18%, mientras que con RandomForestRegressor se alcanzó las siguientes predicciones, para el tiempo de secado: 99.18%, consumo de energía: 99.17% y tasa de germinación: 88.72%. Los modelos generados fueron evaluados mediante la precisión del modelo logrando un porcentaje más alto del 95.47%, por lo tanto, el empleo de estas herramientas permitió obtener resultados óptimos en la predicción de las variables de tiempo de secado, consumo de energía y tasa de germinación, que ayudan a reducir, ahorrar tiempo y recursos.

**Palabras claves:** Modelo, predicción, temperatura, machine learning, secado de semillas.

**TECHNICAL UNIVERSITY OF COTOPAXI.**

**FACULTY OF ENGINEERING SCIENCES AND APPLIED.**

**THEME: "MODEL BASED ON MACHINE LEARNING FOR THE PREDICTION OF VARIABLES USED IN THE SEED DRYING IN A MICROWAVE OVEN".**

**Authors: Cárdenas Quinapaxi Silvia Carolina.**

**Vela Pérez Vanessa Estefanía.**

**ABSTRACT**

Seed drying involves reducing the moisture content at recommended levels, using techniques that do not deteriorate their viability, avoid deterioration, heating and infestation during storage. The present project is based on developing a model based on Machine Learning for predicting the variables used in seed drying, as there is no any model related to this type of research, therefore it is not possible to obtain a good efficiency in drying, which it invests time, economic and material resources.

For the prediction development process, predictive techniques were used, such as: Neural Networks and Decision Trees, in addition, a methodology involving the experimental and prototype method was used. Additionally, the Python programming language was used with the Skylearn tool, this last tool allowed the use of the MLPRegressor neural network, which was calibrated with the following main characteristic; 5 input layers and 5 output layers were used. Likewise for the prediction with decision trees RandomForestRegressor was used, the same one that was calibrated under the main characteristic that in this case was the use of 10 estimators and a random state of 42. MLPRegressor reached the following predictions, for the drying time : 99.18%, energy consumption: 98.17% and germination rate: 78.18%, while with RandomForestRegressor the following predictions were reached, for the drying time: 99.18%, energy consumption: 98.17% and germination rate: 88.72%, The models generated were evaluated by coefficient of variation achieving 95.47%, therefore, the use of these tools allowed ob have optimal results in the prediction of the variables of drying time, energy consumption and germination rate, which help reduce, save time and resources.

**Keywords:** model, prediction, temperature, machine learning, seed drying.

## AVAL DE TRADUCCIÓN



Universidad  
Técnica de  
Cotopaxi

CENTRO DE IDIOMAS

## *AVAL DE TRADUCCIÓN*

En calidad de Docente del Idioma Inglés del Centro de Idiomas de la Universidad Técnica de Cotopaxi; en forma legal **CERTIFICO** que: La traducción del resumen del proyecto de investigación al Idioma Inglés presentado por las señoritas Egresadas de la Carrera de Ingeniería en Informática y Sistemas Computacionales de la Facultad de Ciencias de la Ingeniería y Aplicadas: **SILVIA CAROLINA CÁRDENAS QUINAPAXI Y VELA PÉREZ VANESSA ESTEFANÍA**, cuyo título versa **“MODELO BASADO EN MACHINE LEARNING PARA LA PREDICCIÓN DE VARIABLES EMPLEADAS EN EL SECADO DE SEMILLA EN HORNO MICROONDAS”**, lo realizaron bajo mi supervisión y cumple con una correcta estructura gramatical del Idioma.

Es todo cuanto puedo certificar en honor a la verdad y autorizo a las peticionarias hacer uso del presente certificado de la manera ética que estimaren conveniente.

Latacunga, Julio del 2019.

Atentamente,

  
**Lic. María Fernanda Aguaiza.**  
**DOCENTE CENTRO DE IDIOMAS**  
**C.C. 050345849-9**

x



CENTRO  
DE IDIOMAS

x

## ÍNDICE GENERAL

|  |      |
|--|------|
| DECLARACIÓN DE AUTORÍA .....   | i    |
| APROBACIÓN DEL TRIBUNAL DE TITULACIÓN .....                          | ii   |
| AVAL DEL TUTOR DE PROYECTO DE TITULACIÓN .....                       | iii  |
| AGRADECIMIENTO.....  | iv   |
| DEDICATORIA .....  | v    |
| RESUMEN.....   | viii |
| ABSTRACT.....  | ix   |
| AVAL DE TRADUCCIÓN .....   | x    |
| ÍNDICE GENERAL.....  | xi   |
| ÍNDICE DE TABLAS .....   | xv   |
| ÍNDICE DE FIGURAS.....   | xvi  |
| ÍNDICE DE ANEXOS.....  | xvii |
| 1.- INFORMACIÓN GENERAL.....   | 1    |
| 2. JUSTIFICACIÓN DEL PROYECTO.....                                   | 2    |
| 3. BENEFICIARIOS DEL PROYECTO .....                                  | 3    |
| 4. EL PROBLEMA DE INVESTIGACIÓN .....                                | 3    |
| 4.1. Formulación del problema .....                                  | 3    |
| 5. OBJETIVOS .....   | 4    |
| 5.1. Objetivo general .....  | 4    |
| 5.2. Objetivos específicos.....                                      | 4    |
| 6. ACTIVIDADES Y SISTEMA DE TAREAS EN RELACIÓN A LOS OBJETIVOS ..... | 4    |
| PLANTEADOS.....  | 4    |
| 7. FUNDAMENTACIÓN CIENTÍFICO TÉCNICA.....                            | 5    |
| 7.1. Antecedentes .....  | 5    |
| 7.2. Bases teóricas .....  | 6    |
| 7.2.1. Definición de semilla .....                                   | 6    |
| 7.2.2. Secado de semillas .....                                      | 8    |
| 7.2.3. Parámetros de secado .....                                    | 8    |
| 7.2.3.1. Condiciones del aire ambiente .....                         | 8    |
| 7.2.3.2. Temperatura de secado .....                                 | 8    |
| 7.2.3.3. Humedad inicial del producto .....                          | 9    |
| 7.2.4. Secado de semillas por microondas.....                        | 9    |
| 7.2.5. Métodos de secado de semillas .....                           | 9    |

|   |    |
|---|----|
| 7.2.5.1. Secado natural .....                                       | 10 |
| 7.2.5.3. Secadores por tandas .....                                 | 10 |
| 7.2.6. Tipos de secadores .....                                     | 10 |
| 7.2.6.1. Secadores directos .....                                   | 11 |
| 7.2.6.2. Secadores por lotes.....                                   | 11 |
| 7.2.6.3. Secadores indirectos .....                                 | 11 |
| 7.2.6.4. De acuerdo al método de operación .....                    | 11 |
| Secadores continuos .....   | 11 |
| 7.2.7. Clasificación de los sistemas de secado .....                | 11 |
| 7.2.7.1. Secado a baja temperatura .....                            | 11 |
| 7.2.7.2. Secado a temperatura media.....                            | 12 |
| 7.2.7.3. Secado a alta temperatura.....                             | 12 |
| 7.2.8. Definición de predicción .....                               | 12 |
| 7.2.9. Análisis predictivo.....                                     | 12 |
| 7.2.9.1. Modelos aplicables en el análisis predictivo .....         | 12 |
| Modelos descriptivos.....   | 13 |
| Modelos de decisión.....  | 13 |
| Modelo ensemble o modelado de conjuntos .....                       | 14 |
| 7.2.10. Validación de los modelos .....                             | 14 |
| 7.2.11. Lenguajes de programación .....                             | 14 |
| 7.2.11.1. Lenguaje de máquina .....                                 | 14 |
| 7.2.11.2. Lenguaje de ensamblador.....                              | 15 |
| 7.2.11.3. Lenguaje de alto nivel .....                              | 15 |
| 7.2.11.4. Paradigmas de programación .....                          | 15 |
| 7.2.11.5. Python .....  | 15 |
| Anaconda python.....  | 16 |
| Entornos de desarrollo integrados .....                             | 16 |
| Librerías de python. ....   | 16 |
| Pandas .....  | 16 |
| Numpy.....  | 16 |
| Sklearn.....  | 16 |
| 7.2.12. Machine Learning .....                                      | 17 |
| 7.2.13. Técnicas de Machine Learning.....                           | 17 |
| 7.2.13.1. Árboles de decisión, clasificación y Random Forests ..... | 18 |
| Árboles de decisión .....   | 18 |
| ID3.....  | 18 |
| 7.2.13.2. Árboles de clasificación .....                            | 18 |
| Árboles de clasificación (CART).....                                | 18 |

|  |    |
|--|----|
| 7.2.13.3. Random Forest .....  | 19 |
| 7.2.13.4. Redes Neuronales .....   | 20 |
| Elementos básicos que componen una red neuronal .....                      | 21 |
| Niveles o capas de una red neuronal .....                                  | 21 |
| Arquitecturas neuronales.....  | 22 |
| Según el número de capas .....   | 22 |
| Según el tipo de conexiones .....  | 22 |
| Según el grado de conexión .....   | 23 |
| Estructuras neuronales.....  | 23 |
| Estructura directa.....  | 23 |
| Estructura inversa.....  | 24 |
| Estructura con retardo .....   | 24 |
| Cancelador de ruido .....  | 25 |
| Perceptrón multicapa.....  | 25 |
| 7.2.14. Entrenamiento y prueba del modelo de predicción .....              | 26 |
| 7.2.14.1. Entrenamiento (train) .....                                      | 26 |
| 7.2.14.2. Prueba (test).....   | 26 |
| 7.2.15. Parámetros para evaluar el modelo de predicción.....               | 26 |
| 7.12.15.1. Precisión .....   | 26 |
| 7.12.15.2. Porcentaje de varianza.....                                     | 27 |
| 7.2.16. Parámetro principal para el secado de semillas .....               | 27 |
| 7.2.16.1. Temperatura para el secado.....                                  | 27 |
| 7.2.17. Variables empleadas para la predicción del secado de semillas..... | 27 |
| 7.2.17.1. Tiempo de secado.....  | 27 |
| 7.2.17.2. Consumo de energía .....   | 28 |
| 7.2.17.3. Tasa de germinación.....   | 28 |
| Fases de la germinación .....  | 29 |
| 8. VALIDACIÓN DE LAS PREGUNTAS CIENTÍFICAS O HIPÓTESIS.....                | 29 |
| 8.1. Hipótesis.....  | 29 |
| 9. Metodologías y diseño experimental.....                                 | 29 |
| 9.1. Materiales y métodos .....  | 29 |
| 9.1.1. Tipos de investigación.....   | 29 |
| 9.1.1.1. Investigación bibliográfica .....                                 | 29 |
| 9.1.2. Métodos de investigación.....                                       | 30 |
| 9.1.2.1. Métodos teóricos .....  | 30 |
| Hipotético deductivo .....   | 30 |
| 9.1.2.2. Métodos empíricos .....   | 30 |
| Técnicas de investigación.....   | 30 |

|   |    |
|---|----|
| Entrevista.....   | 30 |
| Observación.....  | 30 |
| 9.1.3. Instrumentos de investigación.....   | 31 |
| 9.1.3.1. Cuestionario de la entrevista .....  | 31 |
| 9.1.4. Métodos específicos .....  | 31 |
| 9.1.4.1. Método experimental.....   | 31 |
| 9.2. Obtención de datos .....   | 32 |
| 9.2.2. Consulta al experto.....   | 33 |
| 9.3. Dataset.....   | 33 |
| 9.4. Técnicas de machine learning utilizadas.....                                   | 34 |
| 9.4.1. Redes Neuronales.....  | 34 |
| 9.4.2. Random Forests.....  | 35 |
| 9.5. Diseño experimental.....   | 35 |
| 9.5.1. Arquitectura del modelo.....   | 35 |
| 9.5.1.1. Algoritmo de predicción de redes neuronales .....                          | 36 |
| 9.5.1.2. Algoritmo de predicción de Random Forest .....                             | 37 |
| 9.5.2. Entrenamiento basado en random forests y redes neuronales .....              | 38 |
| 9.5.3. Evaluación.....  | 38 |
| 10. ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS.....                                     | 39 |
| 10.1. Predicción de las variables correspondientes aplicando redes neuronales ..... | 39 |
| 10.2. Predicción de las variables correspondientes aplicando Random Forests.....    | 41 |
| 10.3. Discusión de los resultados (Predicción).....                                 | 43 |
| 10.4. Evaluación del modelo .....   | 44 |
| 10.4.1. Evaluación del modelo con Redes Neuronales .....                            | 44 |
| 10.4.2. Evaluación del modelo con Random Forests.....                               | 45 |
| 11. IMPACTOS (TÉCNICOS, SOCIALES, AMBIENTALES O ECONÓMICOS).....                    | 46 |
| 11.1. Impacto social .....  | 46 |
| 11.2. Impacto técnico .....   | 46 |
| 12. PRESUPUESTO PARA LA PROPUESTA DEL PROYECTO.....                                 | 47 |
| 12.1. Gastos directos .....   | 47 |
| 12.2. Gastos indirectos .....   | 47 |
| 12.3. Gastos aproximados .....  | 48 |
| 13. CONCLUSIONES Y RECOMENDACIONES.....   | 48 |
| 13.1. Conclusiones .....  | 48 |
| 13.2. Recomendaciones.....  | 49 |
| 14. BIBLIOGRAFÍA.....   | 50 |
| 15. ANEXOS.....   | 53 |

## ÍNDICE DE TABLAS

|  |    |
|--|----|
| <b>Tabla 1.-</b> Tareas de los objetivos planteados. ....  | 4  |
| <b>Tabla 2.-</b> Descripción de las variables utilizadas que se encuentran en el Dataset. ....             | 34 |
| <b>Tabla 3.-</b> Resultados de la predicción de las variables con una temperatura de 35 °C. ....           | 39 |
| <b>Tabla 4.-</b> Resultados de la predicción de las variables con una temperatura de 45 °C. ....           | 40 |
| <b>Tabla 5.-</b> Resultados de la predicción de las variables con una temperatura de 55 °C. ....           | 40 |
| <b>Tabla 6.-</b> Resultados de la predicción de las variables con una temperatura de 35 °C. ....           | 41 |
| <b>Tabla 7.-</b> Resultados de la predicción de las variables con una temperatura de 45 °C. ....           | 42 |
| <b>Tabla 8.-</b> Resultados de la predicción de las variables con una temperatura de 55 °C. ....           | 42 |
| <b>Tabla 9.-</b> Evaluación del modelo de las variables dependientes, aplicando redes neuronales.<br>..... | 45 |
| <b>Tabla 10.-</b> Evaluación del modelo de las variables dependientes, aplicando Random Forests.<br>.....  | 46 |
| <b>Tabla 11.-</b> Gastos directos. ....  | 47 |
| <b>Tabla 12.-</b> Gastos indirectos. ....  | 47 |
| <b>Tabla 13.-</b> Descripción de los gastos. ....  | 48 |

## ÍNDICE DE FIGURAS

|                    |  |    |
|--------------------|--|----|
| <b>Figura 1.-</b>  | Semilla. ....  | 7  |
| <b>Figura 2.-</b>  | Secadora de tandas inclinadas o albercas. ....                                 | 10 |
| <b>Figura 3.-</b>  | Modelo predictivo.....   | 12 |
| <b>Figura 4.-</b>  | Esquema de Random Forest. ....   | 19 |
| <b>Figura 5.-</b>  | Comparación entre una neurona biológica (izquierda) y una artificial (derecha) | 20 |
| <b>Figura 6.-</b>  | Esquema de una red neuronal. ....  | 21 |
| <b>Figura 7.-</b>  | Red neuronal monocapa. ....  | 22 |
| <b>Figura 8.-</b>  | Red neuronal multicapa. ....   | 22 |
| <b>Figura 9.-</b>  | Red neuronal recurrente.....   | 23 |
| <b>Figura 10.-</b> | Esquema de una estructura directa. ....  | 24 |
| <b>Figura 11.-</b> | Esquema de bloques de la estructura inversa. ....                              | 24 |
| <b>Figura 12.-</b> | Esquema de una estructura con retardo. ....                                    | 24 |
| <b>Figura 13.-</b> | Esquema de un cancelador activo de ruido.....                                  | 25 |
| <b>Figura 14.-</b> | Estructura de un perceptrón multicapa. ....                                    | 25 |
| <b>Figura 15.-</b> | Germinación de la semilla. ....  | 28 |
| <b>Figura 16.-</b> | Proceso de obtención de datos.....   | 32 |
| <b>Figura 17.-</b> | Arquitectura del método de predicción.....                                     | 35 |
| <b>Figura 18.-</b> | Algoritmo de predicción del proceso de redes neuronales. ....                  | 36 |
| <b>Figura 19.-</b> | Algoritmo de predicción del proceso de Random Forests.....                     | 37 |

## ÍNDICE DE ANEXOS.

|   |    |
|---|----|
| <b>Anexo 1.-</b> Hoja de vida autor 1 .....   | 54 |
| <b>Anexo 2.-</b> Hoja de vida autor 2 .....   | 55 |
| <b>Anexo 3.-</b> Hoja de vida tutora .....  | 56 |
| <b>Anexo 4.-</b> Certificado .....  | 57 |
| <b>Anexo 5.-</b> Código de la variable tiempo de secado aplicando Redes Neuronales .....    | 58 |
| <b>Anexo 6.-</b> Código de la variable consumo de energía aplicando Redes Neuronales .....  | 59 |
| <b>Anexo 7.-</b> Código de la variable tasa de germinación aplicando Redes Neuronales ..... | 60 |
| <b>Anexo 8.-</b> Código de la variable tiempo de secado aplicando Random Forests .....      | 61 |
| <b>Anexo 9.-</b> Código de la variable consumo de energía aplicando Random Forests .....    | 62 |
| <b>Anexo 10.-</b> Código de la variable tasa de germinación aplicando Random Forests .....  | 63 |

## **1.- INFORMACIÓN GENERAL**

### **Título del Proyecto:**

Modelo basado en Machine Learning para la predicción de variables empleadas en el secado de semilla en horno microondas.

**Fecha de inicio:** Octubre 2018.

**Fecha de finalización:** Julio 2019.

**Lugar de ejecución:** Barrio El Ejido, Parroquia Eloy Alfaro, Cantón Latacunga, Universidad Técnica de Cotopaxi, Proyecto generativo “Secado de semillas empleando horno microondas”.

**Facultad que auspicia:** Ciencias de la Ingeniería y Aplicadas.

### **Carrera que auspicia:**

Ingeniería en Informática y Sistemas Computacionales.

**Proyecto de investigación vinculado:** Proyecto generativo- secado de semillas empleando horno microondas.

### **Equipo de Trabajo:**

**Tutor de Titulación:** Ing. Silvia Jeaneth Bravo Mullo.

**Estudiantes:** Cárdenas Quinapaxi Silvia Carolina.

Vela Pérez Vanessa Estefanía.

**Área de Conocimiento:** Ciencias.

**Sub-área:** Informática.

**Línea de investigación:** Tecnologías de la información y comunicación (TICs) y diseño gráfico.

**Sub líneas de investigación de la Carrera:** Inteligencia artificial e inteligencia de negocios.

## 2. JUSTIFICACIÓN DEL PROYECTO

En la actualidad el sector agrícola requiere de la mejora de métodos para el secado de semillas, los factores más representativos del agente secante que intervienen en este proceso son: la precisión, temperatura, viscosidad y la densidad, estos factores influyen en forma directa en el tiempo y calidad de secado. (Valdivia, 2011)

El presente trabajo de investigación se elaboró con la finalidad de contribuir con el proceso de secado de semilla mediante técnicas de Machine Learning, ya que ello permitirá reducir recursos tales como el tiempo de trabajo, recursos humanos y materiales que se emplean para la realización de esta actividad (Rodríguez, 2018).

Actualmente, la tecnología se ha involucrado en varias áreas del conocimiento, así mismo en el sector agrícola y alimentario, por lo cual su uso resulta en mejoras para la optimización de recursos. (Suntasig, 2017). El aprendizaje de máquina es empleado para realizar predicciones que involucrarían gran consumo de recursos teniendo resultados satisfactorios y de aprovechamiento de recursos. (Zhao-hu & Zhen, 2009). Por lo tanto el presente proyecto emplea técnicas de Machine Learning para la predicción de variables tales como Redes Neuronales y Random Forests, haciendo que el proceso de secado de semillas sea mucho más rápido y con una significativa reducción de costos. (Zhao-hu & Zhen, 2009).

Es importante mencionar que el modelo desarrollado tendrá un impacto social, porque beneficiará a las industrias involucradas en el secado y tratamiento de semillas. El impacto técnico involucra el uso de las tecnologías para optimizar procesos que resultarían en el empleo de gran cantidad de tiempo para observar resultados. El impacto económico que trata sobre el ahorro de recursos humanos y materiales al emplear herramientas de predicción que reducen los errores al mínimo.

Hay que recalcar que esta investigación beneficiará a las personas del sector agrícola y a los investigadores relacionados al ámbito de inteligencia artificial, ya que mediante esta propuesta podrán obtener una herramienta que muestra resultados verídicos cercanos a la realidad en cuanto al secado de semillas en horno microondas.

Esta investigación está realizada por estudiantes que aplicaron sus conocimientos adquiridos en años de estudio en esta prestigiosa institución, es decir, se cuenta con el recurso humano necesario para desarrollar la propuesta prevista. Por otro lado los recursos tecnológicos utilizados en esta investigación son de distribución gratuita por lo cual esta propuesta puede ser empleada de forma gratuita.

### **3. BENEFICIARIOS DEL PROYECTO**

Este proyecto va a beneficiar a las personas del proyecto generativo de “Secado de semillas empleando horno microondas”; así también al sector agrícola y a los investigadores que emplean inteligencia artificial, ya que los resultados que se obtengan permitirán ahorrar tiempo y recursos.

### **4. EL PROBLEMA DE INVESTIGACIÓN**

El objetivo fundamental del secado es retirar el agua del producto hasta un nivel en donde el desarrollo microbiológico y la deterioración del mismo sean disminuidas, aumentar la calidad de los alimentos, limitar los costos de almacenamiento y de transporte, al igual que darle un valor agregado al producto. (Ayala, 2011)

La humedad y la temperatura son variables de mayor importancia en la conservación de las semillas, es por eso que el secado hace referencia a la reducción del contenido de humedad a niveles recomendados, lo cual utiliza técnicas que no deterioren su viabilidad, eviten el deterioro, calentamiento e infestación durante su almacenamiento. (Correa, Alvarez, Espitia, & Cardona, 2013)

A pesar que el secado en microondas es lo adecuado, ya que resulta ser un proceso mucho más rápido que otros métodos evaluados dependiendo de la potencia usada en el equipo (Suárez, Barrera, & Forero, 2016) , aún se emplea mucho tiempo, recursos y energía, por ende mediante una indagación realizada se pudo constatar en diferentes investigaciones la creación de métodos para la predicción de secado de semillas con diferentes procesos, tales como gel de sílice para muestras de semillas pequeñas y en túnel de secado (Suárez et al., 2016), sin embargo, no existe un modelo que presente resultados óptimos y precisos para conservar un buen rendimiento en el secado, retardando el tiempo planificado. (Casabianca et al., 2015).

Las semillas pueden deteriorarse al estar sujeto a una alta desecación que daña el origen y especie de la semilla (Correa et al., 2013). En el sector agrícola, se emplea recursos como tiempo y presupuesto, los mismos que causan problemas de pérdidas económicas por falta de una herramienta tecnológica, la cual permita mejorar la precisión en el cálculo de predicción de las variables empleadas en el secado de semillas en base a una temperatura determinada (Novoa & Palacios, 2010).

#### **4.1. Formulación del problema**

¿Cómo se puede predecir variables como el tiempo de secado, tasa de germinación y consumo de energía para el secado de semillas en horno microondas?

## 5. OBJETIVOS

### 5.1. Objetivo general

Diseñar un modelo de predicción empleando técnicas de Machine Learning, para la determinación de variables empleadas en el secado de semillas en horno microondas.

### 5.2. Objetivos específicos

- Identificar los aspectos que intervienen en la predicción de secado de semilla mediante el análisis de la literatura.
- Identificar las variables a predecir en el secado de semilla en horno microondas.
- Emplear técnicas de Machine Learning para la creación de un modelo para la predicción de secado de semilla en hornos microondas.
- Validar el modelo desarrollado mediante la simulación y posterior comprobación empleando datos reales.

## 6. ACTIVIDADES Y SISTEMA DE TAREAS EN RELACIÓN A LOS OBJETIVOS PLANTEADOS.

Tabla 1.- Tareas de los objetivos planteados.

| Objetivo  | Actividad (tareas)  | Resultado de la actividad               | Medios de verificación                       |
|---|---|---|--|
| Identificar los aspectos que intervienen en la predicción de secado de semilla mediante el análisis de la literatura. | <ul style="list-style-type: none"> <li>• Investigar datos bibliográficos</li> <li>• Determinar los artículos que serán analizados.</li> <li>• Analizar los algoritmos de predicción.</li> </ul> | Revisión de la literatura.              | Marco teórico del proyecto de investigación. |
| Identificar las variables a predecir en el secado de semilla en horno microondas.                                     | <ul style="list-style-type: none"> <li>• Determinar las variables que se predicen de acuerdo a varios autores.</li> <li>• Evaluar las variables que se necesitan predecir.</li> </ul>           | VARIABLES A PREDICIR.                   | Experimentación del proyecto.                |
|   | <ul style="list-style-type: none"> <li>• Determinar las técnicas de Machine Learning y evaluar</li> </ul>   | Cuadro de la experimentación realizada. | Experimentación del proyecto.                |

|  |  |  |                               |
|--|--|--|-------------------------------|
| Emplear técnicas de Machine Learning para la creación de un modelo para la predicción de secado de semilla en hornos microondas. | las técnicas que mejor se adapten al proyecto.   |  |                               |
| Validar el modelo desarrollado mediante la simulación y posterior comprobación empleando datos reales.                           | <ul style="list-style-type: none"> <li>• Determinar las técnicas de Machine Learning.</li> <li>• Determinar las variables.</li> <li>• Desarrollar la experimentación en varias fases.</li> </ul> | Algoritmo de predicción base del modelo. | Experimentación del proyecto. |

**Fuente:** Equipo de trabajo.

## 7. FUNDAMENTACIÓN CIENTÍFICO TÉCNICA

### 7.1. Antecedentes

El secado es una operación esencial en muchas industrias, tales como la química, agrícola, biotecnológica, alimenticia, de polímeros, cerámica, farmacéutica, pulpa y papel, procesamiento de minerales e industrias de transformación de la madera. A menudo, el secado es la operación final de una serie de operaciones que se desarrollan en el proceso de fabricación de un determinado producto y se realiza inmediatamente antes de su envasado y envío. Sin embargo, en algunos casos, el secado es una parte esencial del proceso de fabricación, como por ejemplo en la fabricación de papel. (Moreno, Maqueda, & Ballesteros, 2017)

El secado es un proceso de transferencia en el cual la humedad es removida y absorbida de la superficie de la semilla en forma de vapor por el aire que la rodea. La temperatura ideal de secado de la semilla varía de 35 a 45 °C, dependiendo de la variedad, de la humedad del grano y condiciones ambientales.

Mientras más alto es el contenido inicial de humedad, más alta es la susceptibilidad de la semilla de ser dañada durante el secado, por eso la semilla de granos básicos con humedad entre 30 y 33% (Madurez fisiológica), es recomendable secarlas a temperaturas menores de 40 °C, cuando se baja la humedad a 20 % o menos, las temperaturas de secado se pueden incrementar

hasta un máximo de 45 °C para lograr bajar la humedad de la semilla a 12-14 %. (Valdivia, 2011)

Existen varios métodos para secar las semillas. Algunos métodos de secado para conservación de germoplasma implican altas temperaturas de aire (35 °C - 45 °C), que afectan negativamente la integridad y viabilidad de la semilla. Los métodos más comunes y seguros son el secado mediante la des-humidificación y el secado con gel de sílice que sirve para secar muestras pequeñas.

El procedimiento consiste básicamente en colocar gel de sílice auto-indicador, azul y seco en un desecador o frasco de vidrio con sello hermético. El peso del gel de sílice utilizado debe ser igual al de las semillas para lograr un secado eficiente. (Correa et al., 2013)

El principio del secado, es el de disminuir la disponibilidad del agua para las reacciones enzimáticas y de crecimiento microbiano mediante la eliminación del agua libre de los productos alimenticios. Otro de los objetivos del proceso de secado es la reducción del volumen y peso global en la elaboración de alimentos útiles. (Franco & Malavé, 2013)

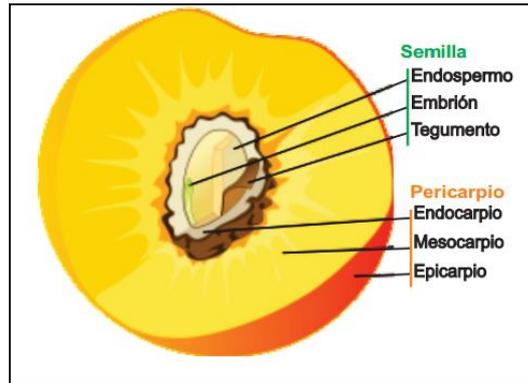
## **7.2. Bases teóricas**

### **7.2.1. Definición de semilla**

Es muy importante destacar que el órgano primordial de cualquier planta terrestre o acuática es la semilla, la misma que cumple funciones importantes para el desarrollo y crecimiento de las plantas. En el medio ambiente, la semilla es un alimento elemental para muchos animales, de igual forma, en la producción agrícola la semilla es fundamental para todo ser humano, lo cual el alimento que se extrae está formado por semillas directa o indirectamente, que sirven del mismo modo como alimento para diferentes animales domésticos. Las semillas se pueden almacenarse por largo tiempo, afirmando así, la conservación de especies y variedades de plantas. (Doria, 2010)

La semilla como se muestra en la Figura 1, es primordial para el origen de una nueva planta, es decir, la semilla se produce por la fecundación del óvulo de una planta gimnosperma, (semilla que no está encerrada dentro del fruto), o a su vez de una planta angiosperma (semilla que está encerrada dentro del fruto y produce flores). Entonces se puede decir que la semilla, es un óvulo fecundado y maduro, la cual contiene un embrión que viene a ser una planta en estado de vida latente. El embrión está conformado por sustancias de reserva y endosperma, la misma que al desarrollarse da lugar al fruto. (Parees, 2008)

**Figura 1.-** Semilla.



**Fuente:** Tomado de (Parees, 2008)

Las semillas, deben ser procesadas para que puedan expulsar las impurezas, para de este modo se puedan clasificar por el tamaño y poder sembrarlas, de igual forma deben destacar la calidad por medio de la separación de las semillas que se encuentren inutilizables o deterioradas, para así aplicar las sustancias del tratamiento sanitario. (Doria, 2010)

#### **7.2.1.1. Calidad de la semilla**

La calidad de la semilla puede dividirse en cuatro condiciones básicas que son: genética, fisiológica, sanitaria y física, es decir que con la presencia de estas cuatro cualidades en estado óptimo permite que la semilla se encuentre en una máxima calidad. (Parees, 2008)

- **Calidad genética:** Se representa por semillas de la misma variedad, de igual forma por plantas que vienen de semillas que tienen las mismas características. Este tipo de calidad especifica el potencial de rendimiento y la calidad productiva. (Parees, 2008)
- **Calidad fisiológica:** Capacidad que tiene la semilla para germinar y dar un nuevo origen a plantas uniformes y con propiedades que determina la actividad y desempeño en el crecimiento de la semilla, esta calidad se enfoca en el porcentaje de germinación y a las propiedades de la semilla. (Parees, 2008)
- **Calidad sanitaria:** Está representada por la desaparición de enfermedades transmisibles por la semilla, es decir la semilla no debe ser portadora de agentes patógenos que generen este tipo de enfermedades. (Parees, 2008)
- **Calidad física:** Se refiere a todas las características que poseen las semillas, es decir la forma, tamaño, color, y la presencia de cualquier cuerpo extraño, como por ejemplo: malezas, materiales inertes, plagas y enfermedades. (Parees, 2008)

### **7.2.2. Secado de semillas**

El secado constituye uno de los métodos que permite separar un líquido de un sólido en general, entendemos por secado la separación de la humedad de los sólidos (o de los líquidos), por evaporación en una corriente gaseosa. (Suntasig, 2017)

El secado es un procedimiento en el que el agua se elimina para impedir o reducir el crecimiento de microorganismos perjudiciales, así como de distintas reacciones químicas. El Departamento de Agricultura de Estados Unidos especifica como producto deshidratado el que no tiene más del 2.5% de agua (base seca), mientras que el alimento seco es todo el producto alimenticio que fue expuesto a un proceso de expulsión de agua y que contiene más del 2.5% de agua (base seca). (Franco & Malavé, 2013)

Cuando culminan los procesos para la extracción y el secado de las semillas, es muy importante conservarlas en las mejores condiciones, para de este modo asegurar el poder germinativo y otros parámetros de calidad. (Doria, 2010)

El secado de semillas, ayuda a la conservación de las proteínas y a dar estabilidad a los productos a secar y evitar el crecimiento de hongos y bacterias, sencillamente el secado permite obtener un producto más estable. Por otro lado hay que tener en cuenta que la operación de secado, es una operación de transferencia de masa de contacto gas- sólido donde la humedad contenida en el sólido se transporta por la evaporación hacia la fase gaseosa, en base a la diferencia, entre la presión de vapor ejercida por el sólido húmedo y la presión parcial de vapor de la corriente gaseosa, es decir cuando las 2 presiones se igualan, se dice que el sólido y el gas está en total equilibrio y el proceso de secado termina. (García, 2013)

### **7.2.3. Parámetros de secado**

Los parámetros que hay que tomar en cuenta para el secado se detallan a continuación, los mismos que permiten establecer el equipamiento adecuado para las distintas condiciones determinadas de secado.

#### **7.2.3.1. Condiciones del aire ambiente**

Establece la cantidad de energía que se requiere para obtener la temperatura de secado, es decir cuando la temperatura ambiente disminuya, mayor será la cantidad de energía que se necesita para calentar el aire, lo cual precisa un mayor valor del secado. (Gan & Sandoval, 2003)

#### **7.2.3.2. Temperatura de secado**

Es el parámetro que mayor prioridad y relevancia tiene al momento de realizar el secado, lo cual interviene en la tasa, en la eficiencia de secado y en la calidad del producto final que se obtenga. (Gan & Sandoval, 2003)

### **7.2.3.3. Humedad inicial del producto**

La humedad inicial es un parámetro que también influye en el secado, por ende, cuando más alto es el contenido de humedad, mayor es la cantidad de agua evaporada por unidad de energía, es decir se emplea un gran porcentaje de energía disponible. (Gan & Sandoval, 2003)

### **7.2.4. Secado de semillas por microondas**

Los microondas se producen por un dispositivo electrónico que se le conoce como magnetrón, la energía que posee el microondas se manifiesta hacia abajo en un conducto rectangular que se lo conoce como guía de la onda o sencillamente se transmiten a través de una ventana transparente de polipropileno hacia la cámara de secado. (Martinez & Mora, 2013)

El secado asistido por microondas ofrece ventajas sobre los secadores convencionales de aire caliente, tales como: los tiempos de secado más cortos, el aumento de la eficiencia energética y el potencial de la reducción del tamaño de los equipos de secado requeridos. (Alvarado, 2017)

El secado por microondas, consiste en que el flujo de aire permita la extracción continua del disolvente evaporado, es por eso que la radiación pasa a través de la ventana de polipropileno hacia la cámara de secado, en donde se extrae el líquido de las semillas. El calor generado con este proceso, elimina la humedad y el vapor producido se expulsa con el flujo del aire, por ende cuando el secado está casi completo, la intensidad del campo de radiación aumenta, ya que los sólidos secos no absorben fácilmente como el agua. (Martinez & Mora, 2013)

El secado asistido por microondas resulta ser un proceso mucho más rápido que los otros métodos evaluados y se registran tiempos de secado entre 15 y 40 min dependiendo de la potencia usada en el equipo. Por otra parte, el secado asistido por microondas permite disminuir significativamente los tiempos del proceso, pero el aumento descontrolado de temperatura en microondas domésticos y la poca disponibilidad tecnológica de equipos microondas para secado industrial. (Suárez et al., 2016)

### **7.2.5. Métodos de secado de semillas**

En el procesado de alimentos se puede utilizar tanto el secado directo como el indirecto. En los secadores indirectos el calor es transmitido hacia el alimento en placas metálicas calientes del recipiente y por contacto directo entre las partículas calientes y frías del alimento. Los secadores directos utilizan gas caliente, normalmente aire, el cual pasa a través del alimento. El calentamiento en los secaderos directos se ve favorecido, y además es más uniforme que en los secaderos indirectos. Los elementos clave para cualquier operación de secado de alimentos que deben ser tenidos en cuenta son la actividad de agua, la temperatura de transición vítrea, los mecanismos y teorías de deshidratación y los cambios físicos y químicos. (Franco & Malavé, 2013)

Existen diferentes métodos de secado para disminuir el contenido de humedad de la semilla, los mismos que se ajustan a las necesidades determinadas que pueden ser económicas o técnicas como se menciona a continuación:

#### 7.2.5.1. Secado natural

Se refiere al secado en el que el movimiento del aire es el principal complemento para este método es decir, se debe a la acción de los vientos y a la radiación solar, la misma que permite evaporar la humedad que se localiza en la parte interior de cualquier tipo de semilla. (Valdivia, 2011)

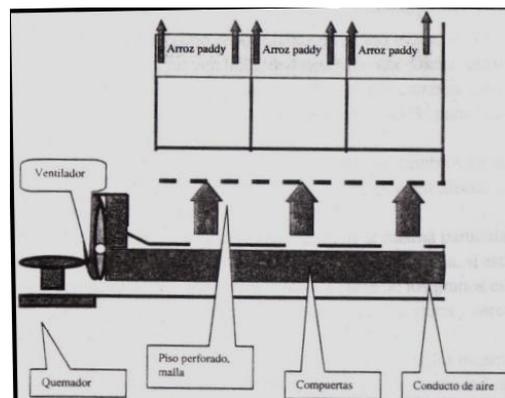
#### 7.2.5.2. Secado artificial

Básicamente se trata de utilizar las propiedades físicas del aire, para de este modo extender su velocidad y cambios de temperatura, para la reducción de la humedad en los lotes de semillas. Esto se consigue con tres técnicas: secamiento estacionario, secamiento continuo, y secamiento intermitente. (Bonilla, 2014)

#### 7.2.5.3. Secadores por tandas

Este tipo de secadoras son conocidas como albercas, como se muestra en la Figura 2, las cuales utilizan mallas en las que se colocan las semillas, para que, de este modo se sequen las capas de los granos con su respectivo proceso de secado, por lo cual son apropiadas para secar cantidades pequeñas, es decir mientras menos es la semilla disminuirá el tiempo de secado.

**Figura 2.-** Secadora de tandas inclinadas o albercas.



**Fuente:** Tomado de (Chippe, 2008)

Esta técnica también se encarga de secar completamente la humedad de la semilla, ya que mediante ello se realiza un estado de secado de forma paulatina y a su vez se obtiene un proceso de secado óptimo dentro del mismo para de esa manera obtener resultados apropiados del secado de la semilla. (Casabianca et al., 2015)

#### 7.2.6. Tipos de secadores

Los secadores pueden clasificarse en diferentes tipos como se muestra a continuación:

### **7.2.6.1. Secadores directos**

La transferencia de calor para la desecación, se consigue por contacto directo entre los sólidos húmedos y los calientes, es por eso que el líquido vaporizado se arrastra con el medio de desecación, es decir con los gases calientes, este tipo de secadores se los llaman también secadores por convección. (García, 2013)

### **7.2.6.2. Secadores por lotes**

Este tipo de secadores se diseña para aplicar con un tamaño específico de lote de alimentación húmeda, para los ciclos de tiempo dado, en estos secadores las condiciones del contenido de humedad y la temperatura varían simultáneamente en cualquier punto determinado. (García, 2013)

### **7.2.6.3. Secadores indirectos**

En estos secadores el calor de la desecación se traslada al sólido húmedo a través de una barrera de retención, en este caso el líquido vaporizado se aparta independientemente del medio de calentamiento, es por eso que la velocidad de desecación requiere del contacto que se plantee entre el material humedecido y las superficies calientes. A los secadores indirectos se los conoce también como secadores por conducción o de contacto. (García, 2013)

### **7.2.6.4. De acuerdo al método de operación**

Se puede clasificar en secadores continuos y discontinuos.

#### **Secadores continuos**

Estos tipos de secadores, se refieren a que secan en menor tiempo, grandes cantidades de cualquier producto de una manera continua. (Novoa & Palacios, 2010)

#### **Secadores discontinuos**

Este tipo de secadores se utilizan para el secado en pequeña y mediana escala, es por eso que la semilla permanece en una cámara hasta que sea secado. (Novoa & Palacios, 2010)

### **7.2.7. Clasificación de los sistemas de secado**

Existe una clasificación de los sistemas de secado sobre la base de la temperatura que se considera necesaria, los mismos que se menciona a continuación:

#### **7.2.7.1. Secado a baja temperatura**

Es aquella que se emplea a una baja temperatura (5 a 8 °C) o a su vez se puede realizar con aire natural. El objetivo es secar la semilla antes de que se deteriore, es decir, si las condiciones climáticas no son las apropiadas se necesitará varios días para empelar este proceso, dependiendo de la humedad inicial que tuvo la semilla. (Bartosik & Rodríguez, 2006)

### 7.2.7.2. Secado a temperatura media

Este secado permite aplicar distintas temperaturas según el uso que se desee dar a la semilla, sin embargo es recomendable una temperatura de 60 °C para la mayoría de semillas y a su vez se lo puede realizar en un lugar o secadora de alta capacidad. (Bartosik & Rodríguez, 2006)

### 7.2.7.3. Secado a alta temperatura

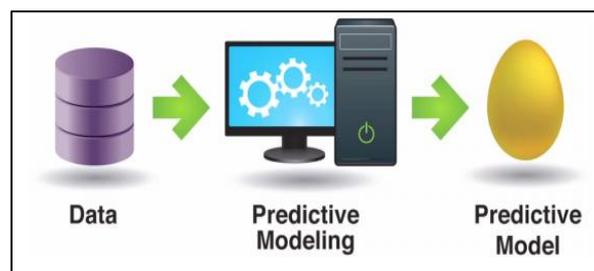
Esta temperatura sobrepasa los 60 ° C, ya que se emplea gran cantidad de flujos de aire en 10° Celsius o un aumento sobre la temperatura ambiente, por ende esta cantidad distingue este tipo de proceso del secado a bajas temperaturas. (Bartosik & Rodríguez, 2006)

### 7.2.8. Definición de predicción

Las predicciones permite identificar las similitudes entre las variables a través de métodos de aprendizaje computacionales y a su vez establecer cálculos adecuados según el proceso de predicción que se esté aplicando, para lo cual se asemejan modelos de procedimientos que permiten la creación de un modelo predictivo. (Espino, 2017)

En la Figura 3, se puede observar un modelo predictivo, el cual se puede utilizar para predecir qué posibilidades existe para que una persona reaccione de una manera establecida en función de los datos obtenidos, es decir elige si desea comprar, cambiar o contratar algún tipo de servicio. Una vez que se haya introducido los datos que requiere la persona y se aplique el modelo predictivo correspondiente, se conseguirá la calificación que muestre la probabilidad estudiada por el modelo. (Espino, 2017)

**Figura 3.-** Modelo predictivo.



**Fuente:** Tomado de (Espino, 2017)

### 7.2.9. Análisis predictivo

El análisis predictivo es idóneo para crear conocimiento sobre los usuarios, es decir no trabaja con la información del individuo en específico, por ende permite identificar pautas de comportamiento que se genera en un conjunto de individuos. (Espino, 2017)

#### 7.2.9.1. Modelos aplicables en el análisis predictivo

Se utiliza el término de análisis predictivo cuando se necesita referirse a los modelos predictivos, lo cual efectúa calificaciones mediante distintos tipos de modelos ya sean

predictivos y pronósticos. Generalmente se está empleando más dicho término, por ende hace referencia a todo lo que corresponde con la disciplina analítica, como el modelado descriptivo o el modelado decisivo, es por eso que esta disciplina contiene un profundo análisis de datos y son ampliamente usadas en la toma de decisiones. (Espino, 2017)

### **Modelos predictivos**

Los modelos predictivos tienen como objeto evaluar la probabilidad de un elemento semejante en un prototipo diferente, para de esta manera demostrar un procedimiento concreto. Este tipo de modelos son utilizados en muchas áreas, como el marketing, donde investigan modelos de datos confidenciales para expresar interrogaciones sobre el desempeño del cliente, o a su vez para modelos de detección de fraude. Los modelos predictivos frecuentemente ejecutan procesos de datos durante las transacciones, es decir evalúan la inseguridad o la oportunidad de un usuario o transacción específica, de forma que contribuyan con conocimientos al momento de tomar una decisión. El análisis predictivo establece un modelo estadístico que emplea los datos existentes para predecir los datos que no se dispone. Para la construcción del modelo predictivo se lo realiza mediante modelos disponibles con propiedades destacadas y un procedimiento conocido, por lo cual se denomina como un conjunto de entrenamiento. (Espino, 2017)

### **Modelos descriptivos**

Este tipo de modelos, permiten medir las relaciones de los datos de manera que es empleada frecuentemente para clasificar clientes o contactos en grupos. Los modelos descriptivos a diferencia de los modelos predictivos, se focaliza en predecir el comportamiento del cliente en especial, en cambio los modelos descriptivos determinan distintas relaciones entre los clientes y los productos. (Espino, 2017)

### **Modelos de decisión**

Especifican la relación de todos los componentes de una decisión, es decir los datos conocidos, teniendo como objetivo predecir los resultados de una decisión la cual intervengan un sinnúmero de variables. Los modelos de decisión son utilizados para modelar una decisión que se elige una sola vez, de igual forma para modelar una guía de toma de decisiones repetible la misma que se puede utilizar una y otra vez, un ejemplo de este tipo de modelos es el árbol de decisión. (Espino, 2017)

### **Modelo ensemble o modelado de conjuntos**

Consiste en la aplicación del modelado predictivo para unir dos o más modelos y después esquematizar los resultados obtenidos en una única puntuación o propagación para optimizar la precisión. Este tipo de modelo considera las predicciones de los dos tipos de modelos caso por caso, en algunas ocasiones puede dar más credibilidad a un modelo que a otro o viceversa, al aplicar este modelo se capacita para predecir qué casos son puntos débiles para cada modelo que lo conforma.

Un ejemplo del modelado de conjuntos, es el modelo de Random Forest (bosque aleatorio), un bosque aleatorio combina árboles de decisión que permite examinar distintos datos de muestra, evaluar diversos factores o variables usuales de peso de manera diferente. (Espino, 2017)

#### **7.2.10. Validación de los modelos**

Es muy importante la comprobación del modelo que se consiguió, es por eso que se realiza la validación correspondiente del mismo, por lo cual lo primero que se realiza es dividir el conjunto de datos que se tiene. Por consiguiente, se dispone de un grupo de datos en donde se creará el modelo, el mismo que comprende dos terceras partes de la muestra denominado *training set*, que se refiere al conjunto de entrenamiento, por otro lado, la tercera parte que viene hacer la última se utiliza para validar el modelo y se denomina *test set*, que hace referencia al conjunto de test.

#### **7.2.11. Lenguajes de programación**

Es importante mencionar que un programa son órdenes o instrucciones que solucionan un problema determinado basado en un lenguaje de programación. Un lenguaje de programación tiene su propia sintaxis, de igual forma se puede decir que es un grupo de reglas, notaciones, símbolos y/o caracteres, lo cual el proceso de datos y estructuras en la computadora es manifestado generalmente por un programador. (Williams, 2010)

El lenguaje de programación consiste en explicar de forma apropiada sobre qué datos se debe manipular en el computador y a su vez, cómo estos datos van a ser guardados o entregados y qué gestiones se debe presentar ante varias circunstancias. Por medio de ello el lenguaje pretende estar aproximadamente al lenguaje humano o natural. (Brueghel, 2008)

##### **7.2.11.1. Lenguaje de máquina**

Este lenguaje es originario de todo computador, es por eso que genera cadenas amplias de unos y ceros, lo cual representa un documento de texto, una imagen, un correo electrónico, etc. (Zapata, 2006)

### 7.2.11.2. Lenguaje de ensamblador

Es el lenguaje que se inventó para manejar el hardware, es decir su estructura corresponde a la identificación total de cada uno de los elementos del ordenador y de esta manera representarlos a través de una serie de registros o campos en la memoria. El conocimiento para este tipo de lenguaje debe ser alto, es decir conocer cada registro que se encuentre almacenado para evitar borrarlo o modificarlo. (Zapata, 2006)

### 7.2.11.3. Lenguaje de alto nivel

Este tipo de lenguajes utilizan términos que están relacionados a las estructuras gramaticales de los lenguajes que utiliza el ser humano, es decir existen palabras y gramática que son empleadas usualmente por las distintas lenguas de habla. (Zapata, 2006)

### 7.2.11.4. Paradigmas de programación

Los paradigmas de programación constituyen tres categorías de los lenguajes de programación como son: imperativos, funcionales y lógicos.

- **Lenguajes imperativos:** Es un conjunto de instrucciones que hace referencia al manejo de la información digital presente en memoria y la manera que se debe almacenar o enviar la información hacia los dispositivos. Algunos ejemplos de este tipo de lenguajes son: Fortran, Algol, C# entre otros.
- **Lenguajes funcionales (o aplicativos):** Es el proceso al que se aplica una función recursiva a un valor de su dominio, para de esta manera obtener el valor respectivo del resultado.
- **Lenguajes declarativos (lógicos):** Se refiere a una especificación de la relación que se busca calcular. (Ureña, 2011)

### 7.2.11.5. Python

Python es un lenguaje de programación interpretado de alto nivel y multiplataforma, fue desarrollado por Guido van Rossum que es un computólogo holandés que en la actualidad trabaja para google. Python es un lenguaje muy sencillo de aprender y de entender, todos los archivos creados en python tienen la extensión .py (Corcuera, 2010)

Python, se trata de un lenguaje de alto nivel, interpretado, interactivo y de propósito general, ya que su diseño hace referencia a una sintaxis limpia y de buena legibilidad y de igual forma está adecuado para la computación científica. Python puede trabajar de dos formas distintas, es decir a través de consola cuando se desea realizar operaciones inmediatas o mediante la ejecución de *scripts* o *guiones de órdenes*, lo cual corresponde a la escritura del código Python que se ejecuta a través del intérprete. (Aranda, 2017)

### **Anaconda python**

Es una distribución que posee el núcleo básico de Python y un conjunto de módulos, de igual manera incluye la consola IPython, que tiene la característica principal de autocompletado, es decir permite teclear las principales letras de una orden y aparece las órdenes que estén disponibles y también incluye el entorno IPython Notebook, que hace referencia a la utilización de un navegador web como interfaz y que compone un entorno de computación que mezcla la edición de texto con el uso de una consola. (Aranda, 2017)

### **Entornos de desarrollo integrados**

Los denominados IDE son programas que ayudan al desarrollo del código, incluyendo un editor de código fuente, el mismo que está acompañado de una consola o herramientas de compilación automática y a menudo también de algún tipo de complemento de depuración o listado de variables. En el caso de Python se puede identificar distintos entornos IDE, los cuales son: IDLE, Stani's Python Editor, Eric IDE, NinJa IDE, Spyder, entre otros, sin embargo Spyder aparece instalado con la distribución ANACONDA, la misma que se puede ejecutar mediante *Anaconda Navigator*. (Aranda, 2017)

### **Librerías de python.**

#### **Pandas**

Es una biblioteca escrita como una extensión de NumPy, la misma que sirve para la manipulación y análisis de datos para el lenguaje de programación Python. En sí, brinda estructuras de datos y operaciones para el manejo de tablas numéricas y series temporales. (Rodríguez, 2018)

#### **Numpy**

Es una extensión open source de Python, la misma que añade mayor soporte para vectores y matrices, por ende constituye una biblioteca de funciones matemáticas de alto nivel para operar con esos vectores o matrices. (Rodríguez, 2018)

#### **Sklearn**

Es una librería que comprende algoritmos de aprendizaje automático de diferentes tipos, como son: clasificación, regresión y agrupación, las mismas que interactúan con las bibliotecas numéricas y científicas de Python. Además es una librería potente y completa al momento de utilizarla en aplicaciones de técnicas de Machine Learning en Python. (Utrera, 2017)

### 7.2.12. Machine Learning

Machine Learning es una rama de la IA y se especifica como el conjunto de conocimientos que logran descubrir modelos automáticamente en un conjunto de datos y utilizar para predecir datos futuros, por lo cual lleva a cabo diferentes ideas de decisiones en ambientes de incertidumbre. (Management Solutions, 2018)

A continuación se detalla algunas aplicaciones de Machine Learning:

- **Predicciones:** Se guarda información de manera digital, que se requiere para la toma de decisiones de una empresa que también se le conoce como Business Intelligence, para lo cual los datos se convierten en predicciones por medio de la utilización de algoritmos, los mismos que ayudan a las empresas a conocer si va a aumentar las ganancias, para así predecir lucros o pérdidas en un futuro. (Rodríguez, 2018)
- **Detección de intrusos:** Los sistemas de IA aprenden el comportamiento de los usuarios autorizados, es por eso que cuando este comportamiento cambia, el sistema elige si se trata de una amenaza y emplea bloqueos necesarios, estos sistemas son conocidos como “detectores de anomalías” (Rodríguez, 2018)
- **Antivirus:** En la actualidad, los antivirus poseen un algoritmo con características esenciales y principales de los malwares que atacan a cualquier archivo ocasionando un problema casi intratable, lo que reduce la complejidad de detectarlos. (Rodríguez, 2018)
- **Clasificación de texto:** Son algoritmos capaces de clasificar texto en base a su contenido, de igual forma, se puede examinar comentarios de redes sociales con fines de marketing. (Rodríguez, 2018)
- **Lingüística computacional:** Se une la lingüística, informática e inteligencia artificial, lo cual se encarga de perfeccionar las interacciones que hay entre las máquinas y los lenguajes humanos. La traducción automática de idiomas, sistemas de recuperación de información y la producción automática de textos, son las aplicaciones que más se conocen. (Rodríguez, 2018)
- **Deep learning:** Es un área de mayor desarrollo, la misma que utiliza algoritmos multinivel para generar distintos niveles de abstracción de la información. Algunos ejemplos de aplicaciones son: el reconocimiento de imágenes, voz, predicción de tendencias digitales, entre otros. (Rodríguez, 2018)

### 7.2.13. Técnicas de Machine Learning

Las técnicas de Machine Learning, consisten en mejorar el proceso de evaluación de modelos, no solo con relación al aumento del poder predictivo por medio de nuevas metodologías y

técnicas de selección de variables, sino también en el avance de la eficiencia de los procesos a través de la automatización.(Management Solutions, 2018)

### **7.2.13.1. Árboles de decisión, clasificación y Random Forests**

#### **Árboles de decisión**

Los árboles de decisión, son una sucesión ordenada de preguntas, las mismas que, la siguiente pregunta requiere de la respuesta a la pregunta actual. Las preguntas son formuladas referentes a las variables que precisan cada elemento, lo que conlleva a que se le asigne a cada una, una determinada clase, dicho procedimiento con las preguntas correspondientes se les representa de forma natural mediante un árbol. (Hidalgo, 2014)

El árbol de decisión, está conformado por un nodo que se conoce como nodo raíz, el mismo que está conectado continuamente con los demás nodos hasta lograr los nodos hoja. Por otro lado, a cada nodo interno se le concede una de las preguntas de la secuencia mientras que al nodo hoja se le establece una etiqueta de clase. (Hidalgo, 2014)

#### **ID3**

Algoritmo que está orientado al uso de variables categóricas, en el caso de que existan variables reales, lo primero que se realiza, es agrupar primero los valores en intervalos para después establecerlas como variables categóricas, es decir sin relación de orden. Una de sus características es que, se considera como variables permitidas para la división del nodo, aquellas que no se ha utilizado, es por eso que el árbol generado por este método, tiene una gran dimensión en los datos. (Hidalgo, 2014)

### **7.2.13.2. Árboles de clasificación**

Los árboles de clasificación agregan un enfoque de clasificación supervisada, esta idea aparece de la estructura de un árbol, el mismo que está conformado de raíz, nodos, ramas y hojas, este tipo de árboles se construye a partir de nodos, los mismos que representan círculos y ramas las mismas que son representadas por los segmentos que conectan los nodos. Este tipo de árbol se empieza con el nodo inicial, el mismo que se le conoce como nodo raíz, se extiende hacia abajo y habitualmente se dibuja de izquierda a derecha. Los nodos que se encuentran en los extremos se los conoce como nodos. (Medina & Ñique, 2017)

#### **Árboles de clasificación (CART)**

Consiste en la construcción de un árbol principal a partir de particiones binarias que se efectúan de una manera sucesiva en los datos, mediante el particionamiento denominado medida de impureza, el mismo que permite establecer la calidad del nodo en la construcción del modelo

de árbol de clasificación, el mismo que establecerá el grado de homogeneidad que existe entre los grupos clasificados según la variable dependiente. (Alarcón, 2017)

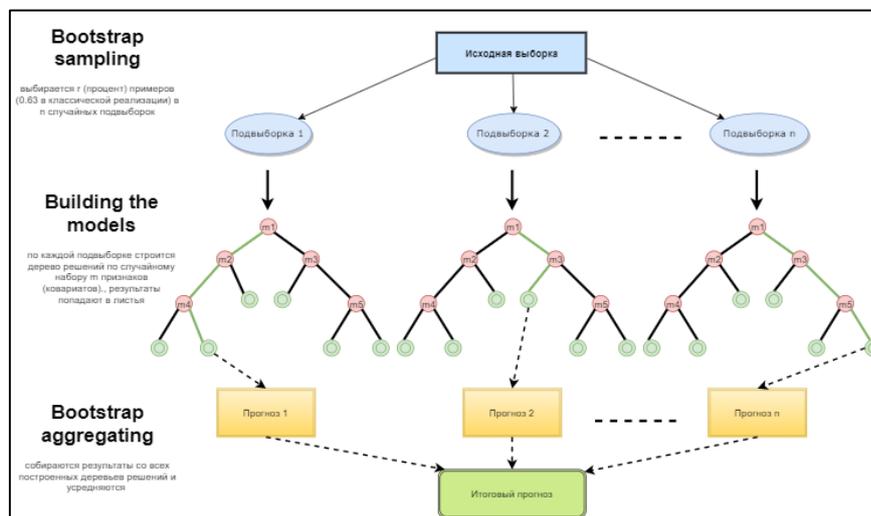
### 7.2.13.3. Random Forest

Random Forests es una técnica de agregación desarrollada por Leo Breiman, la misma que mejora la precisión en una clasificación mediante la incorporación de aleatoriedad en la construcción de cada clasificador individual. (Castrillejo, da Silva, & Illanes, 2010)

Random Forests, trata de la construcción de una serie de árboles predictores, con un subconjunto de los datos, para de este modo promediar los resultados obtenidos, esta técnica es proporcionada para datos con un sinnúmero de distintas características. (Utrera, 2017)

Random Forest (Bosque al azar), se presenta como un clasificador que permite realizar una colección independiente de los clasificadores introducidos en los árboles, ya que cada uno de ellos aporta una variedad en particular. Esta técnica consiste en la combinación de árboles predictivos el cual funciona con una colección de árboles correlacionados y los promedian según cada árbol dependiendo de los valores aleatorios de la muestra de manera independiente con la distribución de los árboles en el bosque, como se muestra en la Figura 4. (Medina & Ñique, 2017)

**Figura 4.-** Esquema de Random Forest.



**Fuente:** Tomado de (Medina & Ñique, 2017)

En distintas ocasiones, la forma de llevar a cabo esta técnica es a partir de un conjunto de números aleatorios como se plantea a continuación:

- **Bagging:** Cada clasificador se construye a partir de una muestra bootstrap, la misma que se genera empleando una cantidad de números aleatorios como elementos tenga el conjunto de entrenamiento. (Hidalgo, 2014)

- **Random Subspace:** Cada clasificador base utiliza únicamente un subconjunto de atributos que se seleccionó aleatoriamente del total de variables, es por eso que cada clasificador se restringe a un subespacio aleatorio de atributos. (Hidalgo, 2014)
- **Técnicas output flipping y class switching:** Se basa en el manejo aleatorio de las etiquetas de clase, por ende se necesita de un número aleatorio para seleccionar datos, cuyas etiquetas sean cambiadas. (Hidalgo, 2014)
- **Randomization:** Introduce aleatoriedad en el algoritmo de aprendizaje, empleando conjuntos de clasificadores con árboles de decisión, en los que el valor de corte es seleccionado aleatoriamente entre los mejores cortes posibles. (Hidalgo, 2014)

#### 7.2.13.4. Redes Neuronales

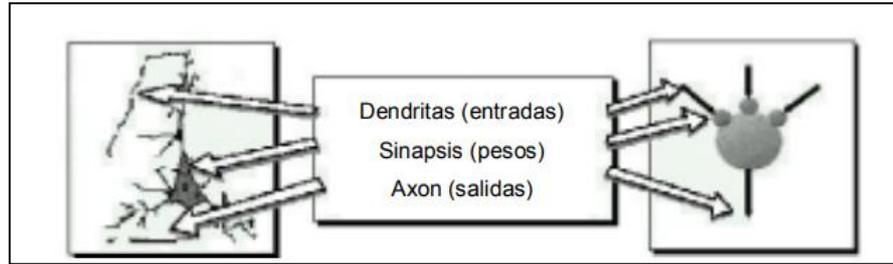
Las redes neuronales son métodos de modelado no lineal sofisticadas que son aptas de modelar funciones complejas, pueden emplearse a problemas de predicción, clasificación o a un control extenso de campos de distintas áreas como por ejemplo de finanzas, psicología cognitiva, medicina, ingeniería y la física. (Espino, 2017)

Las redes neuronales tratan sobre el funcionamiento de la neurona biológica y del funcionamiento de la sinapsis humana. Este tipo de redes son un modelo de aprendizaje y de procesamiento automático, es decir se refiere a un sistema de interconexión entre neuronas que colaboran entre sí para promover un estímulo de salida. (Gutiérrez, 2016)

Las redes neuronales se emplean cuando se desconocen de la naturaleza exacta de la relación entre los valores de entrada y salida, una característica fundamental de las redes neuronales es que aprenden la relación que tienen los valores de entrada y salida a través del entrenamiento. (Espino, 2017)

Para conocer acerca de las redes neuronales hay que conocer acerca de las neuronas y de su utilidad para este tipo de red. En la Figura 5, se realiza la comparación de una neurona biológica con una neurona artificial, por ende se puede observar similitudes entre ambas, es decir tienen entradas, utilizan pesos y generan salidas. (Matich, 2001)

**Figura 5.-** Comparación entre una neurona biológica (izquierda) y una artificial (derecha)

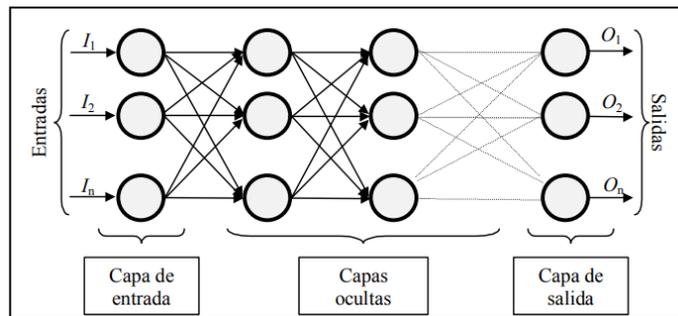


Fuente: Tomado de (Matich, 2001)

### Elementos básicos que componen una red neuronal

En la Figura 6, se muestra un esquema de una red neuronal, la misma que está compuesta por neuronas interconectadas y arregladas en tres capas la misma que puede variar, es decir pueden aumentar el número de capas. Los datos ingresan a través de la “capa de entrada”, los mismos que pasan por medio de la “capa oculta” y salen por la “capa de salida”, hay que tener en cuenta que la capa oculta puede estar formada por varias capas.

Figura 6.- Esquema de una red neuronal.



Fuente: Tomado de (Matich, 2001)

### Niveles o capas de una red neuronal

La distribución de neuronas dentro de la red, se efectúa formando niveles o capas con un número específico de las neuronas en cada una de ellas, es por eso que se puede identificar tres tipos de capas como son:

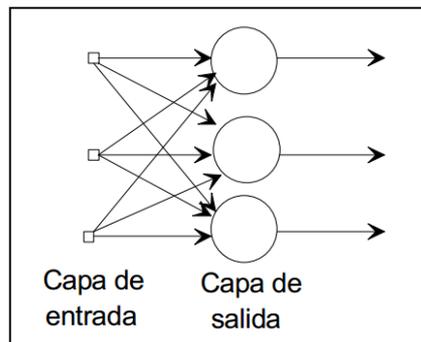
- **De entrada:** Esta capa recibe directamente toda la información de las fuentes externas de la red.
- **Ocultas:** Son internas a la red, lo cual no tienen ninguna relación directa con el medio exterior, el número de niveles ocultos puede estar entre 0 o cualquier número indefinido. Las neuronas de las capas ocultas pueden estar interconectadas de diversas formas, lo cual establece junto con su número, las diferentes topologías de redes neuronales.
- **De salidas:** Transfieren la información de la red hacia el exterior. (Matich, 2001)

## Arquitecturas neuronales

### Según el número de capas

- **Redes neuronales monocapas:** Es la red neuronal más sencilla, como se muestra en la Figura 7, porque posee una capa de neuronas que muestran las entradas a una capa de neuronas de salida en la cual se efectúan distintos cálculos. Un ejemplo típico de este tipo de redes es las memorias asociativas, en la figura se puede observar una estructura de este tipo de red. (Soria, Martín, & Serrano, 2009)

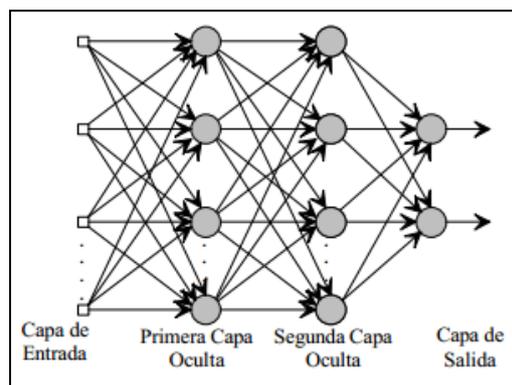
Figura 7.- Red neuronal monocapa.



Fuente: Tomado de (Soria et al., 2009)

- **Redes neuronales multicapa:** Este tipo de redes está formado por un conjunto de capas intermedias de entrada y salida, como se muestra en la Figura 8, es decir capas ocultas, de igual forma puede estar total o parcialmente conectada. (Soria et al., 2009)

Figura 8.- Red neuronal multicapa.



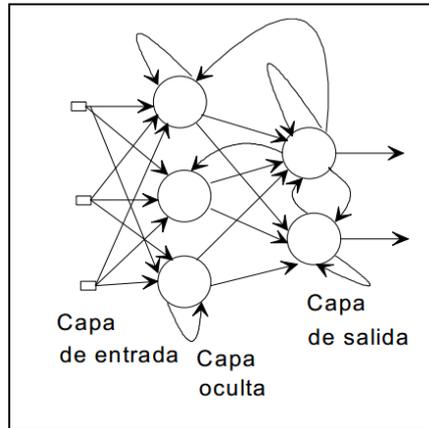
Fuente: Tomado de (Soria et al., 2009)

### Según el tipo de conexiones

- **Redes neuronales no recurrentes:** Este tipo de red la propagación de las señales se realiza simplemente en un sentido, es por eso que no existe la posibilidad de realimentaciones, evidentemente estas estructuras no tienen memoria. (Soria et al., 2009)

- **Redes neuronales recurrentes:** Esta red se caracteriza por la presencia de lazos de realimentación, como se muestra en la Figura 9, estos lazos pueden ser entre neuronas de distintas capas, neuronas de la misma capa o simplemente entre una misma neurona.

**Figura 9.-** Red neuronal recurrente.



**Fuente:** Tomado de (Soria et al., 2009)

### Según el grado de conexión

- **Redes neuronales totalmente conectadas:** En este tipo de redes las neuronas de una capa se localizan conectadas con las neuronas de la capa siguiente (redes no recurrentes) o con las neuronas de la anterior (redes recurrentes). (Soria et al., 2009)
- **Redes parcialmente conectadas:** Este tipo de redes se caracteriza porque no se da la conexión completa entre las neuronas de distintas capas. (Soria et al., 2009)

Esta conexión se puede llevar a cabo de distintas formas siendo las más comunes las estructuras en paralelo y las jerárquicas. En las estructuras en paralelo hacen referencia para obtener la salida, mientras que en las jerárquicas actúan como elementos centrales en la salida final de la red. (Soria et al., 2009)

### Estructuras neuronales

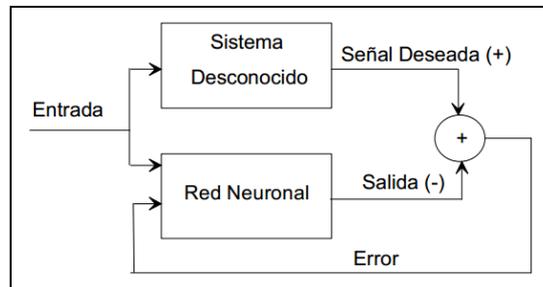
Las redes neuronales se pueden utilizar en distintas estructuras según la aplicación a la que está destinado el sistema, es por eso que dependiendo de la disposición de la red neuronal se podrá obtener las siguientes estructuras: (Soria et al., 2009)

#### Estructura directa

En la Figura 10, se muestra el esquema de bloques, que significa que el sistema en un principio es desconocido y que con la red neuronal poseen las mismas entradas, para así obtener el mínimo error (objetivo de la red neuronal) cuando la salida de la red neuronal y la señal que se necesita sean iguales. Este tipo de estructura tiene como objetivo la modelización de funciones

de transferencia de sistemas, que en un principio, no se conoce nada, sin embargo existe la posibilidad de estimularlos con una entrada definida y de este modo poder conocer su salida. (Soria et al., 2009)

**Figura 10.-** Esquema de una estructura directa.

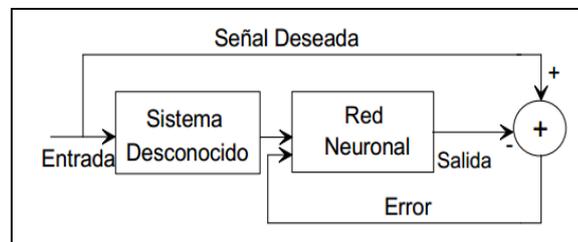


**Fuente:** Tomado de (Soria et al., 2009)

### Estructura inversa

En la Figura 11, se muestra el esquema de una estructura inversa, la misma que el mínimo error en esta estructura se conseguirá cuando la salida de la red neuronal sea la entrada al sistema desconocido lo cual hace referencia que la función de transferencia de la red neuronal sea la inversa del sistema desconocido. (Soria et al., 2009)

**Figura 11.-** Esquema de bloques de la estructura inversa.

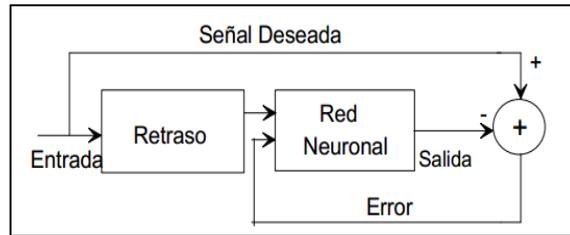


**Fuente:** Tomado de (Soria et al., 2009)

### Estructura con retardo

En la Figura 12, se observa el esquema de una estructura con retardo, la misma que tiene como objetivo disminuir la diferencia entre la señal que se desea, (señal de entrada al instante) y la salida de la red neuronal, la cual será un valor establecido con los valores de dicha señal. Este tipo de estructura se puede aplicar en problemas de predicción y de control. (Soria et al., 2009)

**Figura 12.-** Esquema de una estructura con retardo.

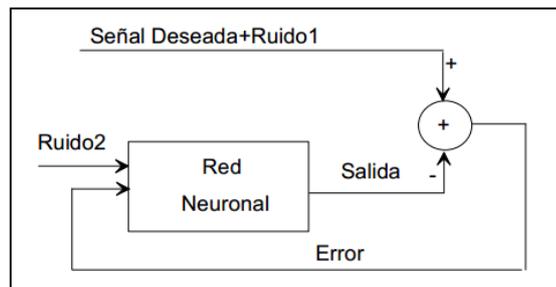


**Fuente:** Tomado de (Soria et al., 2009)

### Cancelador de ruido

En la Figura 13, se muestra el esquema de una red neuronal cuando se emplea como un cancelador de ruido, el objetivo principal de esta estructura lo asume el algoritmo de la red neuronal, es decir lo que realiza este algoritmo es disminuir el error cuadrático entre la salida de la red neuronal y la señal que se toma como referencia.

**Figura 13.-** Esquema de un cancelador activo de ruido.



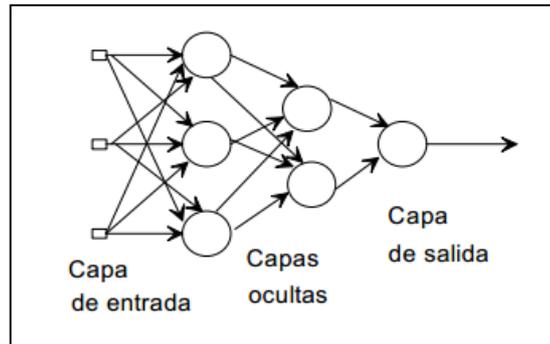
**Fuente:** Tomado de (Soria et al., 2009)

### Perceptrón multicapa

El perceptrón multicapa (MultiLayer Perceptron, MLP), es una red que está constituida por una capa de entrada, un número arbitrario de capas ocultas y una capa de salida, dicha estructura se muestra en la Figura 14. (Soria et al., 2009)

Las neuronas ocultas o de salida toman una entrada de las neuronas de la capa previa, es decir de las conexiones hacia atrás, pero no existen conexiones laterales entre las neuronas dentro de cada capa. La capa de entrada está formada por un conjunto de neuronas, las mismas que corresponden a las variables independientes que se necesita representar, la capa de salida corresponde a la variable de respuesta, por ende es una variable categórica. (Longoni, Porcel, López, & Dapozo, 2010)

**Figura 14.-** Estructura de un perceptrón multicapa.



**Fuente:** Tomado de (Soria et al., 2009)

Se puede establecer cuatro características fundamentales del perceptrón multicapa que son las siguientes:

1. Es una estructura altamente no lineal.
2. Presenta tolerancia a fallos.
3. El sistema es capaz de establecer una relación entre dos tipos de conjuntos de datos.
4. Existe la posibilidad de emplear una implementación de hardware. (Soria et al., 2009)

#### **7.2.14. Entrenamiento y prueba del modelo de predicción**

El entrenamiento y prueba consiste en dividir de manera aleatoria la base de datos en dos partes, una de ellas está conformada por los datos que se utilizarán para entrenar el modelo y los datos restantes servirán como casos de prueba para conocer la precisión de los modelos. (Planas, 2018)

##### **7.2.14.1. Entrenamiento (train)**

Corresponde al conjunto de datos que se va a utilizar para el entrenamiento del modelo, es decir se utilizan unidades de muestra disponibles con atributos y un comportamiento conocido. (Espino, 2017)

##### **7.2.14.2. Prueba (test)**

Hace referencia a los datos que se utiliza para realizar la comprobación del modelo, es por eso que se utiliza una serie de unidades de otra muestra con atributos similares, pero de las cuales no se conoce su comportamiento. (Espino, 2017)

#### **7.2.15. Parámetros para evaluar el modelo de predicción**

##### **7.12.15.1. Precisión**

Método para evaluar la precisión de un modelo al dividir los datos en entrenamiento y prueba, la evaluación de la precisión, muestra la factibilidad que tiene el modelo, es decir si los

resultados que arroja son aproximados al 100%, indica que se puede hacer uso del mismo para distintos experimentos. (Planas, 2018)

#### **7.12.15.2. Porcentaje de varianza**

Se refiere al porcentaje que se obtiene de las predicciones resultantes, es decir la aproximación que tiene el resultado de la predicción con los resultados reales. (Planas, 2018)

#### **7.2.16. Parámetro principal para el secado de semillas**

Para realizar la predicción de las variables para el secado de semilla, hay que tomar en cuenta la temperatura, ya que en base a la misma, se puede predecir el valor de las variables para el secado.

##### **7.2.16.1. Temperatura para el secado**

La temperatura, es el parámetro de mayor flexibilidad que interviene específicamente en la tasa y eficiencia de secado y en la calidad del producto final. Por ende un aumento de la temperatura, significa un menor consumo de energía por unidad de agua evaporada y una mayor tasa de secado, sin embargo una temperatura muy elevada puede originar daños en la semilla. Los flujos de aire y de semillas, establecen la cantidad de agua evaporada en un secador, pero se puede decir y se ha establecido que las temperaturas del aire de secado superiores a 60 ° C, son perjudiciales para la calidad de la semilla. (Novoa & Palacios, 2010)

#### **7.2.17. Variables empleadas para la predicción del secado de semillas**

##### **7.2.17.1. Tiempo de secado**

Para determinar el tiempo de secado, es importante tener en cuenta distintos factores, los más importantes se detallan a continuación:

**Tipo de producto:** Hace referencia a que mayor contenido de agua se emplea mayor tiempo.

**Tamaño de los trozos del producto:** Se refiere a la cantidad del producto, es decir si la cantidad es grande, mayor es el tiempo.

**Temperatura del aire:** Mientras más sea elevada la temperatura, menor es el tiempo.

**Humedad relativa del aire:** Se refiere a que mientras más elevada sea la humedad, mayor es el tiempo.

**Velocidad del aire:** Mientras más sea elevada la velocidad, menor es el tiempo. (Echeverriarza, 2014)

### 7.2.17.2. Consumo de energía

En la actualidad, el consumo de energía es un gran impacto de forma eléctrica como térmica de materiales y de suelo, ya que es utilizado como la energía mundial y a su vez se le conoce como per-cápita. (Castells & Bordas, 2011)

De igual forma, es un componente esencial del análisis global de los procesos de producción industrial, por lo cual diferentes tipos de combustibles y carburantes son utilizados en las industrias así como la magnitud del consumo de los mismos, es por eso que se realiza un análisis clave de la importancia de la utilización de los productos energéticos en el proceso de producción. (Hernández, 2006)

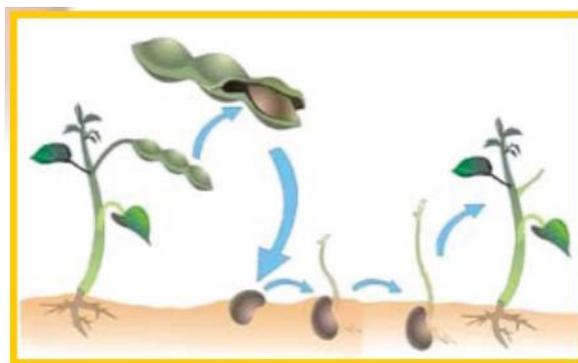
### 7.2.17.3. Tasa de germinación

Hay que destacar que el valor de una semilla se encuentra en el porcentaje de germinación, la misma que es un parámetro fundamental con el cual se establece la cantidad de semillas que se pueden convertir en plantas dentro de un conjunto estipulado de semillas. (Parees, 2008)

La germinación es muy importante, ya que constituye el proceso en el cual se origina las transformaciones metabólicas que se necesitan en el desarrollo de la plántula, en esta fase la absorción del agua se disminuye de manera considerable, la misma que puede llegar a detenerse. (Doria, 2010)

Para que la semilla pueda germinar debe localizar ciertas condiciones favorables, como por ejemplo: que el embrión debe estar completamente formado, maduro, sano y vivo, de igual forma que se encuentre en un lugar óptimo para la germinación, es decir con aire para la respiración, agua para que se hinche y la humedad para reblandecer la cáscara de la semilla, como se muestra en la Figura 15. (Parees, 2008)

**Figura 15.-** Germinación de la semilla.



**Fuente:** Tomado de (Parees, 2008)

## **Fases de la germinación**

Conforma tres etapas sucesivas, las cuales son:

- Absorción de agua por imbibición, la misma que causa su hinchamiento y ruptura final del tegumento exterior que envuelve la semilla.
- Inicio de la actividad enzimática y del metabolismo respiratorio, lo cual asimila las reservas alimenticias en las zonas en crecimiento del embrión.
- Crecimiento y división celular que provoca la emergencia de la radícula y consecutivamente de la plúmula.

Una vez que la semilla haya culminado la fase de la germinación, pasa a una siguiente fase de crecimiento, la misma que tiene que originar una plántula o por el contrario puede morir. (Doria, 2010)

## **8. VALIDACIÓN DE LAS PREGUNTAS CIENTÍFICAS O HIPÓTESIS.**

### **8.1. Hipótesis**

Un modelo basado en Machine Learning para la predicción de variables empleadas en el secado de semilla en horno microondas predecirá datos exactos que pueden ser usados en este contexto.

## **9. Metodologías y diseño experimental**

### **9.1. Materiales y métodos**

#### **9.1.1. Tipos de investigación**

##### **9.1.1.1. Investigación bibliográfica**

Este tipo de investigación, proporciona el conocimiento de investigaciones que ya existen referentes al tema que se va a investigar, de un modo sistemático y través de una extensa indagación de información, conocimientos y distintas técnicas sobre un tema determinado. (Mora, 2010)

En la presente investigación se empleó la investigación bibliográfica, la cual permitió tomar como referencia investigaciones de distintos autores sobre el tema a tratar, para de esta manera sustentar el marco teórico y los resultados de la investigación.

## **9.1.2. Métodos de investigación**

### **9.1.2.1. Métodos teóricos**

#### **Hipotético deductivo**

En este método se hace referencia a la hipótesis, ya que es un punto importante dentro de cualquier investigación, de modo que se pueda comprobar a futuro la veracidad o falsedad del desarrollo de la investigación. (Rodríguez, Pérez, & Alipio, 2017)

En la investigación se utilizó este método, ya que permitió plantear una hipótesis en la que mediante ella se podrá explicar el tema que se está investigando, realizando pruebas que nos ayuden a comprobar si es verdad o es falso el problema a solucionar, comprobando de este modo si se puede obtener posibles soluciones.

### **9.1.2.2. Métodos empíricos**

#### **Técnicas de investigación**

##### **Entrevista**

La entrevista es una técnica que se define como una conversación que se plantea con un fin determinado, generalmente intervienen los investigadores y el sujeto de estudio, de manera que se pueda recopilar los datos más significativos que sean de utilidad para la investigación. (Díaz, Torruco, Martínez, & Varela, 2013)

Mediante esta técnica se consigue información propia de los expertos del tema que se está investigando, es decir para esta investigación, se realizó una entrevista al PhD. Ángel Hernández, lo cual nos permitió conocer detalles más concretos sobre el tema, lo cual fue un aporte muy importante, ya que se obtuvo datos específicos para el desarrollo de la investigación.

##### **Observación**

La observación, es parte fundamental de todo proceso investigativo, por ende es una técnica que consiste básicamente en observar detenidamente el fenómeno, hecho o caso, para de este modo, tomar como referencia toda la información que se pueda obtener y registrarla para su posterior análisis. (Sánchez, 2015)

Mediante esta técnica se pudo establecer un acercamiento con las personas involucradas con el tema a investigar, con el fin de identificar las falencias que existe al momento del secado de semillas, para de este modo dar una posible solución y obtener resultados óptimos y eficientes.

### **9.1.3. Instrumentos de investigación**

#### **9.1.3.1. Cuestionario de la entrevista**

El cuestionario, es un instrumento que se basa en un conjunto de preguntas que se realiza a una persona o a un conjunto de personas sobre un tema o problema específico, del cual se desea conocer, para de este modo obtener más información para la investigación que se está realizando. (Corral, 2010)

En la presente investigación se utilizó el cuestionario para la entrevista con preguntas referentes al tema, lo cual permitieron obtener información necesaria para contribuir en el proceso del marco teórico y conocer de forma precisa y clara sobre el tema a desarrollar. Se realizó las siguientes preguntas a la persona entrevistada:

1. ¿Por qué se enfocó en la realización de la investigación de secado de semilla en microondas?
2. ¿Qué aspectos se consideró en el secado de semilla?
3. ¿Qué se debería predecir en el secado de semillas y cómo se debería hacer?
4. ¿Cuáles son los beneficios del modelo de predicción del secado de semillas para su investigación?

### **9.1.4. Métodos específicos**

#### **9.1.4.1. Método experimental**

Este método permite manipular, ya sea una o diversas variables de estudio, con el fin de poder controlar el aumento o disminución de las mismas. De igual forma se puede decir que un experimento hace referencia al cambio del valor de una variable específica (variable independiente) y observar su efecto en distinta variable (variable dependiente). (Serrano et al., 2008)

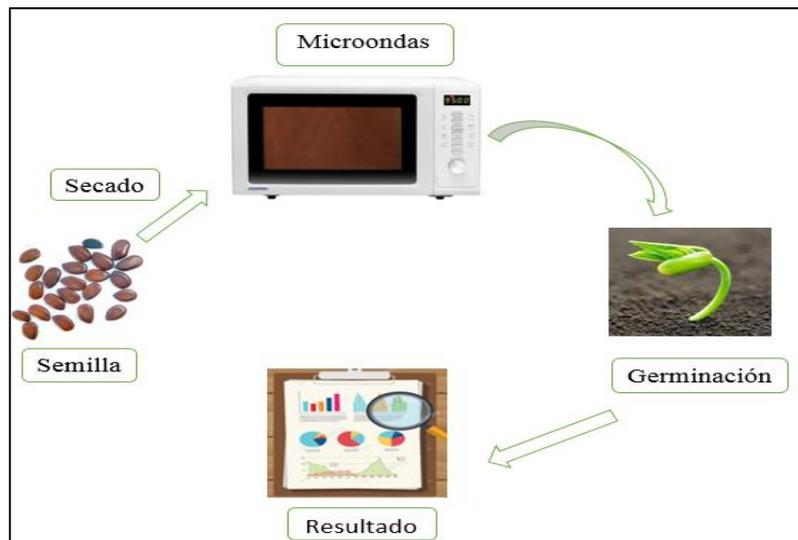
El método seleccionado para el modelo de predicción de variables empleadas en el secado de semillas es el método experimental, lo cual se utilizó un dataset, el mismo que contiene datos investigados. Este tipo de datos fueron calibrados en la herramienta Python, conjuntamente con una librería principal denominada Skylearn y un conjunto de líneas de código. Esta sucesión fue útil para realizar el respectivo proceso de ejecución para las variables empleadas en base a la temperatura aplicada, obteniendo resultados óptimos y precisos. Esta metodología brinda información verídica de la presente investigación, ya que podrá mejorar el tiempo de trabajo y a su vez permitirá seguir realizando una investigación que supere lo investigado.

## 9.2. Obtención de datos

Para la obtención de los datos de la presente investigación, se realizó una secuencia de pasos, los mismos que se muestra en la Figura 16, por otro lado se obtuvo información relevante con la entrevista que se realizó al experto.

### 9.2.1. Gráfica del proceso

**Figura 16.-** Proceso de obtención de datos.



**Fuente:** Equipo de trabajo.

En la Figura 16, se muestra el proceso de secado de las semillas de amaranto, del cual se obtuvieron los datos utilizados en esta investigación. El proceso de secado de estas semillas se lo realizó en un horno microondas doméstico que tiene incorporado un controlador de temperatura, lo cual permitió examinar la temperatura del proceso de secado. En este horno, se colocó una masa de semillas de 100 g, con una humedad inicial de aproximadamente 20% y éstos se secaron a 3 temperaturas (35, 45 y 55 ° C) hasta que se obtuvo una humedad final del 12%. Para cada temperatura se asignó 5 experimentos, en cada experimento el consumo de energía eléctrica se midió usando un analizador de energía. Una vez que las semillas pasan por el proceso de secado se procede a realizar las pruebas de germinación, para lo cual se utilizó una cámara de germinación y una placa Petri en la que se colocó las semillas para cada uno de los experimentos en base a la temperatura.

Concluido todo el proceso que se mencionó anteriormente, se pudo conseguir los datos para las 3 variables que se emplean en el secado de semilla, lo cual todos estos datos conforman el dataset.

### **9.2.2. Consulta al experto**

La presente investigación se realizó con el objetivo de mejorar los recursos de secado de semilla, esta idea surgió a partir del proyecto generativo – secado de semillas empleando horno microondas, lo cual el principal autor es el PhD. Ángel Hernández, mediante una entrevista realizada nos pudo brindar más información acerca del tema para continuar con la investigación.

#### **1. Por qué se enfocó en la realización de secado de semilla en microondas?**

Se realizó esta investigación para determinar los beneficios que se obtiene al igual que el tiempo que se tarda en secar una semilla, considerando distintos factores que son primordiales para el secado.

#### **2. ¿Qué aspectos se consideró en el secado de semilla?**

Se consideró distintos aspectos para el secado de semillas, como: el tiempo de secado, la tasa de germinación y el consumo de energía que son variables que ayudan a determinar y controlar el secado de la misma.

#### **3. ¿Qué se debería predecir en el secado de semillas y cómo se debería hacer?**

Se debe predecir el tiempo que tarda en arrojar datos óptimos de las variables que se consideran para el secado de semillas, para esto se debe diseñar un modelo de predicción basado en técnicas informáticas, el mismo que ayudará a predecir.

#### **4. ¿Cuáles son los beneficios del modelo de predicción del secado de semillas para su investigación?**

Los beneficios que se obtiene con el modelo de predicción son: ahorro de recursos, de tiempo y de energía, al igual que el uso de la tecnología, es decir reduce el período de obtención de datos, ya que mediante todo el proceso que se realiza para el secado de semillas por microondas se tarda de 3 a 6 meses en conseguir resultados eficientes y poder utilizarlos en cualquier experimentación.

### **9.3. Dataset.**

Para la presente investigación se utilizó un dataset en formato CSV, con 5 experimentos realizados por cada temperatura de la semilla de amaranto, lo cual determina datos de las 3 variables, por ende cuenta con 60 datos en total, los mismos que se obtuvieron mediante la consulta al experto y los datos recolectados del secado y germinación de dicha semilla. Este dataset contiene y muestra los valores en base a la temperatura (variable independiente) aplicada del tiempo de secado, consumo de energía y tasa de germinación (variables dependientes), todos los datos son numéricos y de tipo entero, es por eso que son muy importantes para comprobar el modelo de predicción y de esta manera obtener el resultado más

próximo. Hay que recalcar que el dataset contiene pocos datos, ya que para obtener toma un tiempo estimado de 6 meses en realizar la experimentación de los mismos.

En la Tabla 2, se despliega una descripción de las tres variables dependientes (tiempo de secado, el consumo de energía y la tasa de germinación de las semillas) en base a la temperatura (variable independiente), las mismas que se utilizaron para medir en cada uno de los experimentos de secado.

**Tabla 2.-**Descripción de las variables utilizadas que se encuentran en el Dataset.

| <b>Variables</b>                 | <b>Descripción</b>   |
|----------------------------------|--|
| <b>La temperatura de secado.</b> | La temperatura es la base para el secado de la semilla (° C).  |
| <b>El tiempo de secado.</b>      | El tiempo que se requiere para el secado de las semillas, generalmente puede llegar hasta una humedad del 12% (min). |
| <b>El consumo de energía.</b>    | Energía eléctrica consumida por el horno microondas para secar las semillas hasta un 12% de humedad. (Wh).           |
| <b>La tasa de germinación.</b>   | Porcentaje de semillas germinadas (%)  |

**Fuente:** Equipo de trabajo.

#### **9.4. Técnicas de machine learning utilizadas.**

Para esta investigación se utilizó técnicas de Machine Learning que son: Random Forests y Redes Neuronales, las mismas que presentan un mayor rendimiento en la predicción de variables y mejoran el proceso de evaluación de los modelos.

##### **9.4.1. Redes Neuronales**

Para el modelo de predicción, se utilizó la estructura de Perceptrón Multicapa (MLP) de redes neuronales, la misma que se importa mediante la librería sklearn. Esta estructura está formada por una capa de entrada que hace referencia a un conjunto de neuronas que representa las variables independientes, de igual forma de una capa de salida la misma que corresponde a la variable de respuesta. MLP, es el encargado de seleccionar la mejor opción, ya que puede aproximar cualquier función continua en un intervalo hasta el nivel deseado.

Para emplear esta técnica, se utilizó el programa ANACONDA el cual se ejecuta mediante Anaconda Navigator con el entorno de desarrollo Spyder (Python), se importó las librerías (panda, numpy y sklearn) necesarias para la ejecución del mismo, tomando en cuenta que la

librería principal es `sklearn.neuronal_network`, la misma que requiere importar la herramienta `MLPRegressor` que permite predecir los valores aproximados de las variables dependientes.

#### 9.4.2. Random Forests

RandomForest es un método formado por un conjunto de árboles de decisión, que también se le conoce como árboles de clasificación o de regresión. Cada árbol se encuentra constituido por ramas, nodos y hojas, en los nodos se inserta los atributos de los valores que va a depender la función final, los mismos que llegan a las hojas mediante las ramas. Cada nodo hoja almacena los resultados de los datos de entrada.

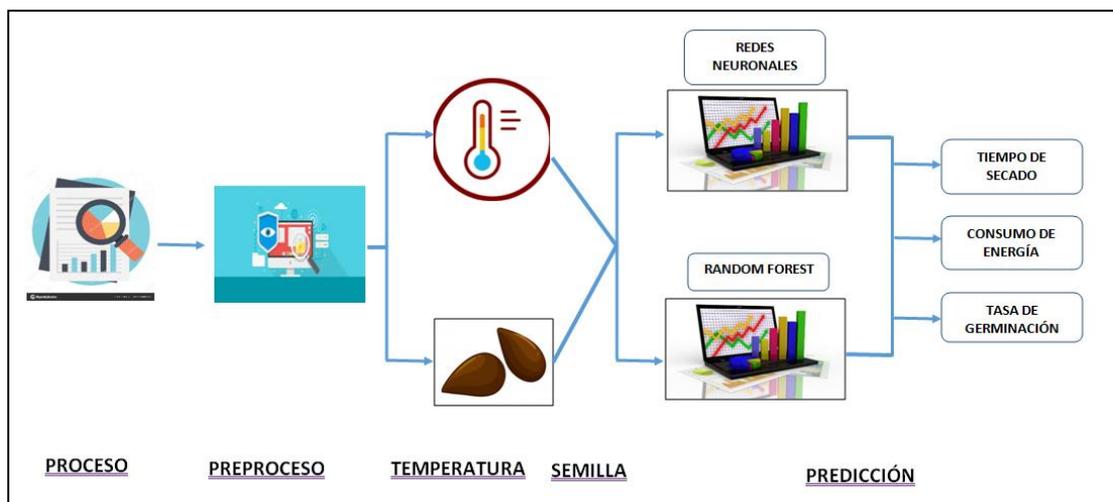
Para emplear esta técnica, se utilizó el programa ANACONDA, el cual se ejecuta mediante Anaconda Navigator con el entorno de desarrollo Spyder (Python), se importó las librerías (panda, numpy y sklearn) necesarias para la ejecución del mismo, tomando en cuenta que la librería principal es `sklearn.ensemble`, la misma que requiere importar la herramienta `RandomForestRegressor` que permite predecir los valores aproximados de las variables dependientes.

### 9.5. Diseño experimental

#### 9.5.1. Arquitectura del modelo

La arquitectura del modelo desarrollado para la predicción del secado de semillas, se muestra en la Figura 17.

**Figura 17.-** Arquitectura del método de predicción.



**Fuente:** Equipo de trabajo.

Como primer paso se ingresa los datos en formato CSV, a continuación en la fase de pre-proceso se descarta todos los datos que no son necesarios y los datos cualitativos se convierten en un formato numérico.

Para el proceso de predicción se utiliza el predictor MLPRegressor para Redes Neuronales y RandomForestRegressor para Random Forests, utilizando la temperatura como entrada. Estos datos se insertan a través del entrenamiento de la Red Neuronal y de Random Forests (ver algoritmo de predicción) y como resultado del proceso de predicción se obtienen el tiempo de secado, consumo de energía y la tasa de germinación.

El algoritmo de predicción se muestra en la sección 9.5.1.1. y 9.5.1.2, mientras que los detalles de los procedimientos de entrenamiento y predicción se muestran en la sección 9.5.2 y finalmente la evaluación se muestra en la sección 9.5.3.

### 9.5.1.1. Algoritmo de predicción de redes neuronales

En la Figura 18, se muestra el algoritmo de predicción aplicando redes neuronales, el mismo que mediante un proceso establecido permite determinar los resultados de las tres variables dependientes. Lo principal de este algoritmo son las capas que se utilizaron y para este caso fueron 5 capas de entrada y 5 capas de salida.

**Figura 18.-** Algoritmo de predicción del proceso de redes neuronales.

```
#Inicio_Proceso
Leer BasedeDatos Td.
Leer Variables (td, Ec, Gr)
x = BasedeDatos (tem)
y = BasedeDatos (td, Ec, Gr)
X = vector x
#Inicio del Bucle
aux1= entrenamiento, prueba (80,20)
aux2= Red_Neuronal (5,5)
aux2 -> frecuencia(aux1)
#fin Bucle
#Calcular la predicción
aux2 -> porcentaje (aux1)
aux -> predicción td
FinProceso
```

**Fuente:** Equipo de trabajo.

Para la predicción, la variable independiente es la temperatura la misma que se almacena en la variable (x). Las variables dependientes son tiempo de secado, consumo de energía y tasa de germinación que se almacenan en la variable (y), como se puede observar en el algoritmo de predicción en la Figura 18.

A la variable X se le da un valor que puede variar dependiendo del conjunto de datos (dataset) y se crea un ciclo repetitivo hasta llegar al valor máximo del porcentaje de la predicción del modelo.

Finalmente se entrena la herramienta de predicción con MLPRegressor (Red Neuronal) asignando el número de capas, en este caso fueron 5 y de igual forma se dividió el modelo en 80% para entrenamiento y 20% para prueba.

### 9.5.1.2. Algoritmo de predicción de Random Forest

En la Figura 19, se muestra el algoritmo de predicción aplicando Random Forests, el mismo que mediante un proceso establecido permite determinar los resultados de las tres variables dependientes. Lo principal de este algoritmo son los estimadores y el estado aleatorio que se utilizaron y para este caso fueron 10 y 42.

**Figura 19.-** Algoritmo de predicción del proceso de Random Forests.

```
#Inicio_Proceso
Leer BasedeDatos Td.
Leer Variables (td, Ec, Gr)
x = BasedeDatos (tem)
y = BasedeDatos (td, Ec, Gr)
X = vector x
#Inicio del Bucle
aux1= entrenamiento, prueba (80,20)
aux2= Random_Forests (10,42)
aux2 -> frecuencia(aux1)
#fin Bucle
#Calculr la predicción
aux2 -> porcentaje (aux1)
aux -> predicción td
#FinProceso
```

**Fuente:** Equipo de trabajo.

Esta técnica es similar a la de red neuronal, por ende la variable independiente es la temperatura, la misma que se almacena en la variable (x). Las variables dependientes son: tiempo de secado, consumo de energía y tasa de germinación, las mismas que se almacenan en la variable (y), como se puede observar en el algoritmo de predicción en la Figura 19.

A la variable (X) se le da un valor que puede variar dependiendo del conjunto de datos (dataset), por lo que se crea un ciclo repetitivo hasta llegar al valor máximo del porcentaje de la predicción del modelo.

Se utilizó el predictor RandomForestRegressor (Random\_Forests) asignando 10 estimadores (número de árboles) y un estado de 42 (variable para evitar variaciones aleatorias). De igual forma se dividió el modelo en un 80% para entrenamiento y 20% para prueba.

### 9.5.2. Entrenamiento basado en random forests y redes neuronales

Para realizar el entrenamiento del modelo de predicción, se empleó Redes Neuronales y Random Forests.

#### Redes neuronales

Se empleó esta técnica con la estructura de MLP (Perceptrón Multicapa), para lo cual se dividió el modelo en 2 partes que son entrenamiento y prueba, utilizando un 80% para entrenar el modelo y un 20% para realizar las respectivas pruebas el mismo que es analizado para obtener un resultado aproximado, de igual forma fue necesario utilizar **MLPRegressor**, que es una herramienta y clasificador para la predicción, el mismo que analiza un conjunto de datos. MLP permite asignar el número de capas, por ende se asignó 5 capas de entrada y 5 capas de salida para obtener una mayor precisión en los resultados, se aplicó **mlr**, que es una técnica estadística para predecir el resultado de una variable de respuesta, **fit** que hace referencia a la frecuencia y rapidez que tiene el entrenamiento de los datos y **score** que presenta el % del modelo de predicción.

La característica principal de redes neuronales, es que se utilizan cuando no se conoce la naturaleza exacta de la relación entre los valores de entrada y de salida, es por eso que las redes neuronales aprenden la relación entre los valores de entrada y salida a través del entrenamiento.

#### Random Forests

Para esta técnica fue necesario utilizar **RandomForestRegressor**, que es un clasificador que tiene alta velocidad de evaluación y realiza procesos de regresión aleatoria de una gran cantidad de datos, de igual forma como en la técnica de redes neuronales, se dividió el modelo en 2 partes que son entrenamiento y prueba, utilizando un 80% para entrenar el modelo y un 20% para realizar las respectivas pruebas el mismo que es analizado para obtener mejores resultados, se aplicó **mlr**, que es una técnica estadística para predecir el resultado de una variable de respuesta, **fit** que es la frecuencia y rapidez que tiene el entrenamiento de los datos y **score** que muestra el % del modelo de predicción. **RandomForestRegressor**, permite asignar el número de estimadores que en este caso es 10 y el estado aleatorio que es 42, por lo que se obtiene una mayor precisión en los resultados.

### 9.5.3. Evaluación

Se aplicó las técnicas de Redes Neuronales y Random Forests para la creación del modelo de predicción, por ende se utilizó distintas librerías que fueron indicadas para el desarrollo del mismo y de esta manera conseguir el valor de la predicción. Los datos que se ingresaron para dicho modelo se encuentran en un dataset en formato CSV, los mismos que se obtuvieron mediante la consulta al experto y los datos recolectados del secado y germinación de dicha

semilla (amaranto). Para obtener los resultados más próximos de la predicción apropiada de las variables dependientes (tiempo de secado, consumo de energía y tasa de germinación), se toma como base a la variable independiente (temperatura).

Para la evaluación del modelo es importante la variable `train_test_split`, para conocer la factibilidad del mismo. Es importante también identificar la precisión del modelo, que se refiere a la exactitud y probabilidad que tiene el mismo y el porcentaje de varianza que hace referencia al puntaje que se obtiene de las predicciones resultantes, es decir si el resultado es cercano a los valores reales.

## 10. ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS

### 10.1. Predicción de las variables correspondientes aplicando redes neuronales

Los resultados de la predicción aplicando la técnica de redes neuronales para las variables de tiempo de secado, consumo de energía y la tasa de germinación en base a las 3 temperaturas aplicadas se muestra desde la Tabla 3 hasta la Tabla 5, tomando en cuenta que se dividió el modelo en 2 partes, para lo cual para el entrenamiento se utilizó el 80% y para la prueba el 20%.

**Tabla 3.-** Resultados de la predicción de las variables con una temperatura de 35 °C.

| <b>TEMPERATURA</b>         | <b>35 °C</b>          |                  |
|----------------------------|-----------------------|------------------|
|                            | <b>PREDICCIÓN (%)</b> | <b>RESULTADO</b> |
| <b>TIEMPO DE SECADO</b>    | <b>99.18</b>          | 330 (min)        |
| <b>CONSUMO DE ENERGÍA</b>  | <b>98.17</b>          | 906.60(Wh)       |
| <b>TASA DE GERMINACIÓN</b> | <b>78.18</b>          | 85.00 (%)        |

**Fuente:** Equipo de trabajo.

Como se muestra en la Tabla 3, aplicando una temperatura de 35 °C, se puede observar las predicciones de las tres variables dependientes, la misma que se obtuvo una predicción máxima de 99.18% para la variable de tiempo de secado, con un resultado de 330 min, un 98.17% para la variable de consumo de energía obteniendo un valor de 906.60 Wh y un 78.18% para la variable de tasa de germinación obteniendo un porcentaje de 85.00%, por ende los resultados obtenidos son aproximados a los resultados reales.

**Tabla 4.-** Resultados de la predicción de las variables con una temperatura de 45 °C.

| <b>TEMPERATURA</b>         | 45 °C                 |                  |
|----------------------------|-----------------------|------------------|
|                            | <b>PREDICCIÓN (%)</b> | <b>RESULTADO</b> |
| <b>TIEMPO DE SECADO</b>    | <b>99.18</b>          | 205 (min)        |
| <b>CONSUMO DE ENERGÍA</b>  | <b>98.17</b>          | 469.66(Wh)       |
| <b>TASA DE GERMINACIÓN</b> | <b>78.18</b>          | 59.36 (%)        |

**Fuente:** Equipo de trabajo.

En la Tabla 4, se muestra que aplicando una temperatura de 45 °C se obtiene las predicciones de las tres variables dependientes, la misma que se obtuvo una predicción máxima de 99.18 % para la variable de tiempo de secado, obteniendo un resultado de 205 min, para la variable de consumo de energía se logró un 98.17% , con un resultado de 469.66 Wh y un 78.18% para la variable de tasa de germinación con un porcentaje de 59.36%, por ende los resultados obtenidos son aproximados a los resultados reales.

**Tabla 5.-** Resultados de la predicción de las variables con una temperatura de 55 °C.

| <b>TEMPERATURA</b>         | 55 °C                 |                  |
|----------------------------|-----------------------|------------------|
|                            | <b>PREDICCIÓN (%)</b> | <b>RESULTADO</b> |
| <b>TIEMPO DE SECADO</b>    | <b>99.18</b>          | 150 (min)        |
| <b>CONSUMO DE ENERGÍA</b>  | <b>98.17</b>          | 320.74 (Wh)      |
| <b>TASA DE GERMINACIÓN</b> | <b>78.18</b>          | 33.73 (%)        |

**Fuente:** Equipo de trabajo.

Finalmente, aplicando una temperatura de 55 °C, como se muestra en la Tabla 5 se puede observar las predicciones de las tres variables dependientes, la misma que se obtuvo una predicción máxima de 99.18% para la variable de tiempo de secado obteniendo un resultado de 150 min, para la variable de consumo de energía se consiguió un 98.17% con un resultado de 320.74 Wh y un 78.18% para la variable de tasa de germinación, con un porcentaje de 33.73%, por ende los resultados obtenidos son aproximados a los resultados reales.

Una vez obtenido los resultados aplicando la técnica de redes neuronales como se puede observar desde la Tabla 3 hasta la Tabla 5, se puede observar una disminución en las variables de tiempo de secado, consumo de energía y tasa de germinación en base a la temperatura aplicada.

Para lo cual, para las tres temperaturas aplicadas (35, 45, 55 °C) la predicción más alta de las variables dependientes fue en la variable de tiempo de secado con un valor de 99.18%, por otro lado para el consumo de energía tuvo una predicción máxima de 98.17% y para la tasa de germinación se obtuvo una predicción de 78.18%, estos resultados se aproximan a los datos reales que se obtuvieron.

### 10.2. Predicción de las variables correspondientes aplicando Random Forests

Los resultados de la predicción aplicando Random Forests, de las variables de tiempo de secado, el consumo de energía y la tasa de germinación, en base a las 3 temperaturas aplicadas, se muestra desde la Tabla 6 hasta la Tabla 8, tomando en cuenta que se dividió el modelo en 2 partes, para lo cual para el entrenamiento se utilizó el 80% y para la prueba el 20%.

**Tabla 6.**-Resultados de la predicción de las variables con una temperatura de 35 °C.

| <b>TEMPERATURA</b>         | <b>35 °C</b>          |                  |
|----------------------------|-----------------------|------------------|
|                            | <b>PREDICCIÓN (%)</b> | <b>RESULTADO</b> |
| <b>TIEMPO DE SECADO</b>    | <b>99.18</b>          | 330 (min)        |
| <b>CONSUMO DE ENERGÍA</b>  | <b>98.17</b>          | 906.38 (Wh)      |
| <b>TASA DE GERMINACIÓN</b> | <b>88.72</b>          | 82.55 (%)        |

**Fuente:** Grupo de trabajo.

Como se muestra en la Tabla 6, se puede observar las predicciones de las tres variables dependientes, la misma que se obtuvo una predicción máxima de 99.18% para la variable de tiempo de secado obteniendo un resultado de 330 min, un 98.17% para la variable de consumo de energía, logrando un resultado de 906.38 Wh y un 88.72% para la variable de tasa de germinación con un porcentaje de 82.55%, por ende los resultados obtenidos son aproximados a los resultados reales.

**Tabla 7.-** Resultados de la predicción de las variables con una temperatura de 45 °C.

| <b>TEMPERATURA</b>         | 45 °C                 |                  |
|----------------------------|-----------------------|------------------|
|                            | <b>PREDICCIÓN (%)</b> | <b>RESULTADO</b> |
| <b>TIEMPO DE SECADO</b>    | <b>99.18</b>          | 204.45(min)      |
| <b>CONSUMO DE ENERGÍA</b>  | <b>98.17</b>          | 472.25 (Wh)      |
| <b>TASA DE GERMINACIÓN</b> | <b>88.72</b>          | 74.80 (%)        |

**Fuente:** Grupo de trabajo.

En la Tabla 7, se muestra las predicciones de las tres variables dependientes, aplicando una temperatura de 45 °C, la misma que se obtuvo una predicción máxima de 99.18% para la variable de tiempo de secado con un resultado de 204.45 min, un 98.17% para la variable de consumo de energía obteniendo un resultado de 472.25 Wh y un 88.72% para la variable de tasa de germinación logrando un porcentaje del 74.80%, por ende los resultados obtenidos son aproximados a los resultados reales.

**Tabla 8.-** Resultados de la predicción de las variables con una temperatura de 55 °C.

| <b>TEMPERATURA</b>         | 55 °C                 |                  |
|----------------------------|-----------------------|------------------|
|                            | <b>PREDICCIÓN (%)</b> | <b>RESULTADO</b> |
| <b>TIEMPO DE SECADO</b>    | <b>99.18</b>          | 149.73 (min).    |
| <b>CONSUMO DE ENERGÍA</b>  | <b>98.17</b>          | 322.87 (Wh)      |
| <b>TASA DE GERMINACIÓN</b> | <b>88.72</b>          | 23.64 (%)        |

**Fuente:** Grupo de trabajo.

Finalmente, como se puede observar en la Tabla 8, aplicando una temperatura de 55 °C, se muestra las predicciones de las tres variables dependientes, la misma que se obtuvo una predicción máxima de 99.18% para la variable de tiempo de secado con un resultado de 149.73 min, un 98.17% para la variable de consumo de energía logrando un resultado de 322.87 Wh y un 88.72% para la variable de tasa de germinación con un porcentaje de 23.64%, por ende los resultados obtenidos son aproximados a los resultados reales.

Una vez obtenido los resultados aplicando la técnica de Random Forests, como se puede observar desde la Tabla 6 hasta la Tabla 8, se puede observar una disminución en las variables de tiempo de secado, consumo de energía y tasa de germinación en base a la temperatura aplicada.

Para lo cual, para las tres temperaturas aplicadas (35, 45, 55 °C), la predicción más alta de las variables dependientes fue en la variable de tiempo de secado con un valor de 99.18%, para la variable de consumo de energía se obtuvo una predicción máxima de 98.17% y para la tasa de germinación se obtuvo una predicción de 88.72%, estos resultados se aproximan a los datos reales que se obtuvieron.

### **10.3. Discusión de los resultados (Predicción)**

Para la técnica de Redes Neuronales, como se puede observar en la Tabla 5, el modelo predice una disminución en el tiempo de secado, es decir con el aumento de la temperatura de 35°C a 55°C y aplicando la predicción más alta que se pudo lograr con el análisis respectivo se obtiene una disminución de 330 min a 150 min.

El modelo realizado en esta investigación, también predice una disminución en el consumo de energía y en la tasa de germinación. Para el consumo de energía, cuando se aumente el valor de la temperatura el resultado es menor, es por eso que se reflejó una disminución aplicando una temperatura de 35 °C a 55 °C, reduciendo de 906.60 Wh a 320.74 Wh respectivamente, sin embargo este resultado también depende de la predicción aplicada, en este caso la predicción óptima fue de 98.17%, tomando en cuenta que este valor se tomó como referencia para las 3 temperaturas aplicadas luego del análisis correspondiente.

Finalmente para la variable de tasa de germinación es la variable que da resultados con una predicción baja, es por eso que después de un análisis determinado, se pudo alcanzar una predicción de 78.18% la misma que se aplicó para las 3 temperaturas, tomando en cuenta que esta predicción es menor a las demás variables (tiempo de secado y consumo de energía), por ende aplicando una temperatura de 35 °C a 55 °C se obtuvo una reducción de 85.00% a 33.73% respectivamente.

Para la técnica de Random Forests como se puede observar en la Tabla 8, el modelo predice una disminución en la variable de tiempo de secado, aplicando un aumento de temperatura de 35°C a 55°C obteniendo una reducción de 330 min a 149.73 min respectivamente y empleando una predicción de 99.18% para todas las temperaturas, la cual se pudo determinar con un análisis respectivo.

De igual forma, el modelo realizado en esta investigación, predice también una disminución en el consumo de energía y en la tasa de germinación. Para el caso del consumo de energía, como se muestra en la Tabla 8, aplicando la misma predicción (98.17%) para las tres temperaturas se obtiene una disminución de 906.38 Wh a 322.87 Wh, es por eso que mientras más alta sea la temperatura menor es el consumo de energía.

Finalmente, al igual que aplicando redes neuronales, la variable de tasa de germinación es la variable que da resultados con una predicción baja a diferencia de las variables de tiempo de secado y de consumo de energía, por ende aplicando una temperatura de 35 °C a 55 °C y con una predicción de 88.72% se obtuvo una reducción en el resultado de 82.55% a 23.64% respectivamente.

Una vez culminado con la aplicación de Redes Neuronales y Random Forest que son técnicas de Machine Learning, se pudo comprobar que los resultados para las variables de tiempo de secado, consumo de energía y tasa de germinación no varían en su totalidad, al igual que en el valor de la predicción, sin embargo se empleó estas técnicas ya que Redes Neuronales trabaja mediante capas (entrada, oculta y salida) y Random Forests emplea árboles de regresión, es por eso que estas técnicas crean distintos modelos para de este modo obtener resultados óptimos y fiables.

#### **10.4. Evaluación del modelo**

Para la evaluación del modelo desarrollado, se aplicó la precisión y el porcentaje de varianza, para lo cual, la precisión hace referencia a la exactitud y probabilidad que muestra el modelo al momento de realizar la predicción con las diferentes variables, es decir muestra si el modelo es factible aplicarlo, por otro lado el porcentaje de varianza, se refiere al puntaje que se obtiene de las predicciones resultantes, es decir la aproximación que tiene el resultado de la predicción con los resultados reales.

##### **10.4.1. Evaluación del modelo con Redes Neuronales**

En la Tabla 9, se muestra los resultados de la predicción del modelo y del porcentaje de varianza obtenidos para cada variable dependiente, aplicando redes neuronales.

**Tabla 9 .-**Evaluación del modelo de las variables dependientes, aplicando redes neuronales.

| <b>VARIABLES</b>            | <b>PRECISIÓN DEL MODELO</b> | <b>PORCENTAJE DE VARIANZA</b> |
|-----------------------------|-----------------------------|-------------------------------|
| <b>TIEMPO DE SECADO.</b>    | 95.47%                      | 95.00%                        |
| <b>CONSUMO DE ENERGÍA.</b>  | 92.69%                      | 93.00%                        |
| <b>TASA DE GERMINACIÓN.</b> | 78.18%                      | 78.00%                        |

**Fuente:** Equipo de trabajo.

Como se muestra en la Tabla 9, la precisión más alta que obtuvo el modelo fue de 95.47% correspondiente al tiempo de secado, para la variable de consumo de energía se obtuvo un 92.69% y para la tasa de germinación la precisión del modelo es baja con un 78.18%, sin embargo hay que recalcar que los datos obtenidos por las tres variables se acerca al 100%, es por eso que el modelo es fiable y se aplicó en este caso. De igual forma se muestra el porcentaje de varianza, para lo cual el puntaje más alto se obtuvo en la variable de tiempo de secado obteniendo un 95.00%, para el consumo de energía se obtuvo un porcentaje de 93.00% y un 78.00% para la tasa de germinación, por lo tanto se puede decir que el porcentaje de varianza de las tres variables se aproxima al 100%, el mismo que hace referencia a la predicción obtenida, mostrando una precisión cercana a los datos reales.

#### **10.4.2. Evaluación del modelo con Random Forests.**

En la Tabla 10, se muestra los resultados de la predicción del modelo y del porcentaje de varianza obtenidos para cada variable dependiente, aplicando Random Forests.

**Tabla 10.-**Evaluación del modelo de las variables dependientes, aplicando Random Forests.

| <b>VARIABLES</b>            | <b>PRECISIÓN DEL MODELO</b> | <b>PORCENTAJE DE VARIANZA</b> |
|-----------------------------|-----------------------------|-------------------------------|
| <b>TIEMPO DE SECADO.</b>    | 95.47%                      | 95.00%                        |
| <b>CONSUMO DE ENERGÍA.</b>  | 92.69%                      | 93.00%                        |
| <b>TASA DE GERMINACIÓN.</b> | 72.91%                      | 73.00%                        |

**Fuente:** Equipo de trabajo.

Después de un análisis correspondiente, tanto aplicando Redes Neuronales y Random Forests, se obtuvo la misma precisión del modelo en las variables de tiempo de secado y consumo de energía, obteniendo un 95.47% para el tiempo de secado y un 92.69% para el consumo de energía, como se muestra en la Tabla 10, para la tasa de germinación la precisión del modelo es baja con un 72.91%, sin embargo hay que recalcar que los datos obtenidos por las tres variables se acerca al 100%, es por eso que el modelo es fiable y se aplicó en este caso. De igual forma aplicando las 2 técnicas antes mencionadas, se pudo obtener el mismo porcentaje de varianza en la variable de tiempo de secado con un 95.00% y en la variable de consumo de energía con un 93.00%, pero para la tasa de germinación el porcentaje es bajo con un 73.00%, recalcando que el porcentaje de varianza de las tres variables se aproxima al 100%, el mismo que hace referencia a la predicción obtenida, mostrando una precisión cercana a los datos reales.

## **11. IMPACTOS (TÉCNICOS, SOCIALES, AMBIENTALES O ECONÓMICOS).**

### **11.1. Impacto social**

El presente proyecto de investigación tuvo un impacto social, porque mediante ello se desarrolló un modelo de predicción de secado de semillas basado en técnicas de Machine Learning, la misma que es de gran ayuda para las personas del sector agrícola, es decir permitió conocer la predicción de las variables empleadas que son: tiempo de secado, consumo de energía y tasa de germinación en base a la temperatura aplicada.

### **11.2. Impacto técnico**

Esta investigación tuvo un impacto técnico, desarrollando una aplicación que permite predecir los valores de las variables dependientes, ya que mediante ello se hace uso de la tecnología

aplicando herramientas informáticas que ayudan a todas las personas del sector agrícola, mejorando el estilo de vida y trabajo de cada uno de ellas.

### 11.3. Impacto económico

Esta investigación tuvo un impacto económico que trata sobre el ahorro de recursos humanos y materiales al emplear herramientas de predicción que reduce errores al mínimo.

## 12. PRESUPUESTO PARA LA PROPUESTA DEL PROYECTO.

### 12.1. Gastos directos

Tabla 11.- Gastos directos.

| DETALLE      | CANTIDAD | V. UNITARIO | V. TOTAL       |
|--------------|----------|-------------|----------------|
| Impresiones. | 450      | \$0,05      | \$22,50        |
| Anillados.   | 3        | \$2,50      | \$7,50         |
| Empastados.  | 2        | \$25        | \$50,00        |
| Copias       | 20       | 0,03        | \$0,60         |
| Flash Memory | 2        | \$7,00      | \$14,00        |
| <b>TOTAL</b> |          |             | <b>\$94,60</b> |

Fuente: Grupo de trabajo

### 12.2. Gastos indirectos

Tabla 12.- Gastos indirectos.

| DETALLE      | CANTIDAD | PERSONAS | V. UNITARIO | V. TOTAL        |
|--------------|----------|----------|-------------|-----------------|
| Transporte   | 48 días  | 2        | \$4,00      | \$384,00        |
| Alimentación | 48 días  | 2        | \$1,25      | \$120,00        |
| <b>TOTAL</b> |          |          |             | <b>\$504,00</b> |

Fuente: Grupo de trabajo

### 12.3. Gastos aproximados

**Tabla 13.-** Descripción de los gastos.

| <b>Descripción</b>                                | <b>Total</b>    |
|---|-----------------|
| Gastos directos.                                  | \$94,60         |
| Gastos indirectos.                                | \$504,00        |
| <b>Total gastos directos + gastos indirectos.</b> | <b>\$598,60</b> |

**Fuente:** Grupo de trabajo

## 13. CONCLUSIONES Y RECOMENDACIONES

### 13.1. Conclusiones

- En este trabajo se ha desarrollado un modelo para la predicción de las variables principales del proceso de secado de semillas mediante microondas, para lo cual se utilizó las técnicas de Machine Learning como: Redes Neuronales y RandomForest.
- El modelo ayuda a predecir las variables de tiempo de secado, consumo de energía y tasa de germinación de las semillas mediante la temperatura aplicada, es decir mientras se aumente el valor de la temperatura menor será el resultado.
- Las predicciones óptimas y máximas que se realizaron con el modelo utilizando las dos técnicas de Machine Learning fueron: para el tiempo de secado: 99.18%, para el consumo de energía: 98.17%, y para la tasa de germinación fue un: 88.72%, estas predicciones se obtuvieron mediante un análisis de resultados en el lenguaje de programación Python.
- El modelo de predicción desarrollado, mediante las dos técnicas aplicadas consiguió resultados adecuados siendo óptimos para cada variable respectiva, sin embargo hay que recalcar que aplicando Random Forests los resultados de la predicción varían constantemente ya que el dataset contiene pocos datos lo que impide que de resultados factibles.
- Se realizó un artículo científico del presente proyecto de investigación, el cual se envió a la revista International Journal of Engineering Research & Technology, el mismo que se encuentra en la fase de evaluación.

### 13.2. Recomendaciones

- Buscar diferentes métodos y técnicas para la creación de nuevos algoritmos que contribuyan en el proceso de predicción de las variables de tiempo de secado, consumo de energía y tasa de germinación, con la finalidad de obtener resultados confiables.
- Emplear otras técnicas de Machine Learning, distintas a las que se utilizaron en esta investigación para de este modo se pueda elegir cual algoritmo es el más apropiado para realizar determinado tipo de predicción y obtener resultados fiables.
- Buscar nuevos métodos para mejorar los resultados obtenidos en esta investigación, para de este modo mejorar el algoritmo desarrollado y poder obtener resultados óptimos.
- Realizar más experimentos con un Dataset que contenga una cantidad significativa de datos, de igual forma aplicar distintas técnicas de Machine Learning para comprobar la fiabilidad del modelo.

## 14. BIBLIOGRAFÍA

- Alarcón, J. (2017). *Modelos de minería de datos: random forest y adaboost, para identificar los factores asociados al uso de las TIC (internet, telefonía Fija y televisión de paga) en los hogares del Perú. 2014.*
- Alvarado, M. B. (2017). Estudio del proceso de secado de fresa usando horno microondas/Study of the stramberry drying process using microwave dryer. *Prospectiva*, 15(1), 29–34. <https://doi.org/10.15665/rp.v15i1.658>
- Aranda, E. (2017). *Índice general.* 1–143.
- Ayala, F. D. (2011). *Tesis de Maestría en Ciencias Biológicas.*
- Bartosik, R., & Rodríguez, J. (2006). *Secado de granos.*
- Bonilla, N. (2014). *Guía Técnica Buenas prácticas de acondicionamiento de semillas de granos básicos ; Infraestructura , y equipamiento Contenido.*
- Brueghel, P. (2008). *Lenguajes de programación Lenguajes de programación.* (1563).
- Casabianca, L. M., Velandia, V. A., Roldan, J. P., Bermudez, N., Rocha, C., & Barrero, I. N. (2015). *Rediseño , Montaje y Estandarización Operativa del Secador Solar de Semillas en la Hacienda La Cosmopolitana.*
- Castells, X., & Bordas, S. (2011). Energía, agua, medioambiente, territorialidad y sostenibilidad. *Diaz de Santos*, 25.
- Castrillejo, A., da Silva, N., & Illanes, G. (2010). *Consistencia de random forests y otros clasificadores promediados.* 1–22. Retrieved from <http://openaccess.uoc.edu/webapps/o2/bitstream/10609/65426/6/rgagoTFM0617memoria.pdf>
- Chippe, F. (2008). *Mejoramiento de una secadora por tandas de una piladora de arroz.*
- Corcuera, P. (2010). *Introducción a la Programación en Python Introducción a la Programación en Python Tabla de contenidos.*
- Corral, Y. (2010). Diseño de cuestionarios para recolección de datos. *Revista Ciencias de La Educación*, 20(36), 17.
- Correa, E. M., Alvarez, S. C., Espitia, M. M., & Cardona, C. E. (2013). Modelos De Secado Y Tolerancia a La Deseccación De Semillas De. *Revista De Ciencias Agrícolas*, 30(302), 20–33.
- Díaz, L., Torruco, U., Martínez, M., & Varela, M. (2013). La entrevista, recurso flexible y dinámico Laura. *Proceedings of 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference, IAEAC 2017*, 2, 1113–1117. <https://doi.org/10.1109/IAEAC.2017.8054186>
- Doria, J. (2010). GENERALIDADES SOBRE LAS SEMILLAS: SU PRODUCCIÓN, CONSERVACIÓN Y ALMACENAMIENTO. *Cultivos Tropicales*, 31(1), 00–00. Retrieved from [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S0258-59362010000100011](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0258-59362010000100011)
- Echeverriarza, M. (2014). *Guía de uso de secadores solares.* Retrieved from <http://www.unesco.org/new/fileadmin/MULTIMEDIA/FIELD/Montevideo/pdf/ED-Guiasecaderosolar.pdf>
- Espino, C. (2017). “Análisis predictivo: técnicas y modelos utilizados y aplicaciones del mismo- herramientas Open Source que permiten su uso.”
- Franco, M. E. B., & Malavé, K. S. E. (2013). *ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL Facultad de Ingeniería en Mecánica y Ciencias de la Producción Previo a la obtención del Título de : INGENIERAS DE ALIMENTOS Presentada por : Carmina Stefania Quinde Fuentes Nancy Viviana Sánchez Lluquín GUAYAQUIL.*
- Gan, A., & Sandoval, N. (2003). Diseño Del Sistema De Control Del Proceso De Secado De Pastas Alimenticias. *Revista Colombiana de Tecnologías de Avanzada*, 2(2), 1–6.

- García, M. del S. (2013). Diseño de un secador de bandejas para el secado de maíz, quinua y amaranto en la Hacienda San Jorge. *Facultad de Ciencias, Bachelor*, 144. Retrieved from <http://dspace.esPOCH.edu.ec/bitstream/123456789/4850/3/96T00332.PDF>
- Gutiérrez, J. (2016). Líneas de investigación en minería de datos en aplicaciones en ciencia e ingeniería: Estado del arte y perspectivas. *Arxiv, Artificial Intelligence (cs.AI)*, 1(1609.05401), 1–17. <https://doi.org/10.1007/s003350010211>
- Hernández, J. (2006). 9. *Consumos energéticos*. 101–111.
- Hidalgo, S. (2014). *Random Forests para detección de fraude en medios de pago*.
- Longoni, M. G., Porcel, E. A., López, M. V., & Dapozo, G. N. (2010). *Modelos de Redes Neuronales Perceptrón Multicapa y de Base Radial para la predicción del rendimiento académico de alumnos universitarios*. 692–701.
- Management Solutions. (2018). *Machine Learning una pieza clave en la transformación de los modelos de negocio*. Retrieved from <https://www.managementsolutions.com/sites/default/files/publicaciones/esp/machine-learning.pdf>
- Martínez, J., & Mora, K. (2013). *Elaboración de un guión y un video de un estudio de la operación unitaria de secado en polvos*". 82.
- Matich, D. J. (2001). Redes Neuronales: Conceptos Básicos y Aplicaciones. *Cátedra: Informática Aplicada a La Ingeniería de Procesos-Orientación I*, 6. Retrieved from [https://www.frro.utn.edu.ar/repositorio/catedras/quimica/5\\_anio/orientadora1/monograia\\_s/matich-redesneuronales.pdf](https://www.frro.utn.edu.ar/repositorio/catedras/quimica/5_anio/orientadora1/monograia_s/matich-redesneuronales.pdf)
- Medina, R., & Ñique, C. (2017). Bosques aleatorios como extensión de los árboles de clasificación con los programas R y Python. *Interfases*, 0(10), 165. <https://doi.org/10.26439/interfases2017.n10.1775>
- Mora, N. (2010). *La investigación Bibliográfica Ideas principales y secundarias*. 1–2. Retrieved from [https://fido.palermo.edu/servicios\\_dyc/blog/docentes/trabajos/17306\\_55962.pdf](https://fido.palermo.edu/servicios_dyc/blog/docentes/trabajos/17306_55962.pdf)
- Moreno, Á. H., Maqueda, R. H., & Ballesteros, I. (2017). *Secado industrial con energía microondas 5*.
- Novoa, W., & Palacios, J. C. (2010). *Diseño de dos sistemas de secado de maíz para el Sector Agrícola del Cantón Ventanas de la Provincia de Los Ríos, Proyecto SENACYT-EPN-PETROCOMERCIAL*. 202.
- Parees, C. (2008). *Boletín Técnico La semilla*. (1).
- Planas, A. (2018). Estudio comparativo de modelos de machine learning para la detección de dianas microARN. *Published by Universitat Oberta de Catalunya*.
- Rodríguez, A., Pérez, J., & Alipio, O. (2017). Métodos científicos de indagación y de construcción del conocimiento. *Revista EAN*, (82), 175–195. <https://doi.org/10.21158/01208160.n82.2017.1647>
- Rodríguez, J. M. (2018). *Aplicación de técnicas de Machine Learning a la detección de ataques*. Retrieved from <http://openaccess.uoc.edu/webapps/o2/bitstream/10609/81126/11/jmrodriguez85TFM0618memoria.pdf>
- Sánchez, M. I. S. (2015). Técnicas de investigación. *Implementación de Guías de Práctica Clínica En El Sistema Nacional de Salud. Manual Metodológico*., 97–112. Retrieved from <https://www.monografias.com/trabajos14/investigacion/investigacion.shtml>
- Serrano, A., García, L., León, I., García, E., Gil, B., & Lea, R. (2008). *MÉTODOS DE INVESTIGACIÓN DE ENFOQUE EXPERIMENTAL Asignatura: Métodos de investigación en Curso: 3º Educación Especial*. Retrieved from <http://www.postgradoune.edu.pe/documentos/Experimental.pdf>
- Soria, E., Martín, J., & Serrano, A. (2009). *REDES NEURONALES ARTIFICIALES*.

- Suárez, L., Barrera, R., & Forero, A. (2016). *Evaluación de alternativas de secado en el proceso de elaboración de harina de lombriz*. 17(1), 55–71.
- Suntasig, M. (2017). “*EVALUACIÓN DE LA VIABILIDAD Y CALIDAD DE LA SEMILLA DE QUINUA (Chenopodium quinoa), TRAS DIFERENTES MÉTODOS DE SECADO.*”
- Ureña, C. (2011). *Lenguajes de Programación Capítulo 1. Introducción*. 59. Retrieved from <http://lsi.ugr.es/~curena/doce/lp/tr-11-12/lp-c01-pres.pdf>
- Utrera, R. G. (2017). *Uso de algoritmos de aprendizaje automático aplicados a bases de datos genéticos*. (July). <https://doi.org/10.13140/RG.2.2.34739.94241>
- Valdivia, R. (2011). *Secamiento de granos o semilla de maíz*. 1–9.
- Williams, J. M. G. . (2010). *Lenguajes*. Retrieved from <http://www.iqcelaya.itc.mx/~vicente/Programacion/Lenguajes.pdf>
- Zapata, C. (2006). *Fundamentos de programación, Guía de autoenseñanza*. 70–74.
- Zhao-hu, L. I. U., & Zhen, W. (2009). *Machine Learning. Inteligencia artificial que está transformando al mundo*.

# 15. ANEXOS

**Anexo 1.- Hoja de vida autor 1****DATOS PERSONALES****Nombres:** Silvia Carolina**Apellidos:** Cárdenas Quinapaxi.**Fecha de Nacimiento:** 15 de Septiembre de 1993.**Edad:** 25 años**Nacionalidad:** Ecuatoriana.**Cedula de Identidad:** 172317065-8**Dirección Domiciliaria:** Aloasí- Barrio Miraflores Alto (Machachi)**Teléfono Domicilio:** 022-309-029**Teléfono Celular:** 0999493483**Estado Civil:** Soltera.**Correo Electrónico:** [silvia.cardenas8@utc.edu.ec](mailto:silvia.cardenas8@utc.edu.ec)**ESTUDIOS PRIMARIOS****Institución:** “Escuela Isabel Yánez”**Dirección:** Cantón Mejía.**ESTUDIOS SECUNDARIOS****Institución:** “Instituto Tecnológico Superior Aloasí”**Dirección:** Av. Víctor Velasco- Parroquia Aloasí- Cantón Mejía.**Título de bachiller:** Técnico en Comercio y Administración.**Especialización:** Aplicaciones informáticas.**ESTUDIO DE TERCER NIVEL****Institución:** Universidad Técnica de Cotopaxi**Carrera:** Ingeniería en Informática y Sistemas Computacionales.**Dirección:** Av. Simón Rodríguez s/n Barrio El Ejido Sector San Felipe.

-----

**FIRMA**

## Anexo 2.- Hoja de vida autor 2

### **DATOS PERSONALES**

**Nombres:** Vanessa Estefanía.

**Apellidos:** Vela Pérez.

**Fecha de Nacimiento:** 05 de Septiembre de 1994.

**Edad:** 24 años.

**Nacionalidad:** Ecuatoriana.

**Cedula de Identidad:** 180466449-6

**Dirección Domiciliaria:** Av. Rumiñahui y Atahualpa (Ambato)

**Teléfono Celular:** 0995584107.

**Estado Civil:** Soltera

**Correo Electrónico:** [vanessa.vela6@utc.edu.ec](mailto:vanessa.vela6@utc.edu.ec)



### **ESTUDIOS PRIMARIOS**

**Institución:** “Unidad Educativa Rodríguez Albornoz”

**Dirección:** Av. Los Shyris y 12 de Octubre. Cantón Ambato.

### **ESTUDIOS SECUNDARIOS**

**Institución:** “Instituto Tecnológico Superior Hispano América”.

**Dirección:** Cantón Ambato

**Título de bachiller:** En comercio y administración.

**Especialización:** Aplicaciones informáticas.

### **ESTUDIO DE TERCER NIVEL**

**Institución:** Universidad Técnica de Cotopaxi

**Carrera:** Ingeniería en Informática y Sistemas Computacionales.

**Dirección:** Av. Simón Rodríguez s/n Barrio El Ejido Sector San Felipe.

-----  
**FIRMA**

**Anexo 3.- Hoja de vida tutora**

**UNIVERSIDAD TÉCNICA DE COTOPAXI**  
**DATOS INFORMATIVOS PERSONAL DOCENTE**

---

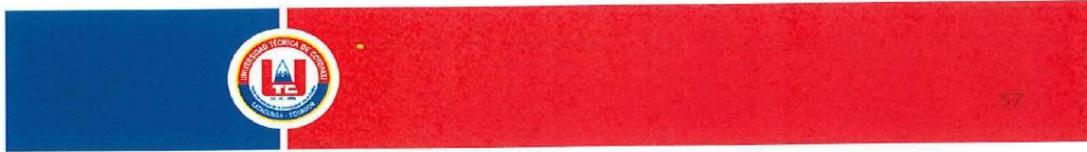
**DATOS PERSONALES****APELLIDOS:** Bravo Mullo**NOMBRES:** Silvia Jeaneth**ESTADO CIVIL:** Casada**CEDULA DE CIUDADANÍA:** 0502437122**NÚMERO DE CARGAS FAMILIARES:** 0**LUGAR Y FECHA DE NACIMIENTO:** Latacunga, 28/11/82**DIRECCIÓN DOMICILIARIA:** Ave. José María Velazco Ibarra y Diego Noboa, 3-39**TELÉFONO CONVENCIONAL:** 032386446    **TELÉFONO CELULAR:** 0984473586**EMAIL INSTITUCIONAL:** silvia.bravom@utc.edu.ec**TIPO DE DISCAPACIDAD:****# DE CARNET CONADIS:****ESTUDIOS REALIZADOS Y TÍTULOS OBTENIDOS**

| <b>NIVEL</b>  | <b>TITULO OBTENIDO</b>   | <b>CÓDIGO DEL REGISTRO CONESUP O SENESCYT</b> | <b>FECHA DE REGISTRO</b> |
|---------------|--|---|--------------------------|
| <b>TERCER</b> | Ingeniera en Informática y Sistemas Computacionales, Universidad Técnica de Cotopaxi | 1020-07-781174                                | 2007-09-07               |
| <b>CUARTO</b> | Master en Tecnologías para la Gestión y Práctica Docente                             | 1027-12-6027930                               | 2012-10-11               |
| <b>CUARTO</b> | Doctor en Ingeniería de Sistemas e Informática                                       | En curso                                      |                          |

**HISTORIAL PROFESIONAL****UNIDAD ADMINISTRATIVA O ACADÉMICA EN LA QUE LABORA:** CIYA**ÁREA DEL CONOCIMIENTO EN LA CUAL SE DESEMPEÑA:** INFORMÁTICA**FECHA DE INGRESO A LA UTC:** 05/05/2008

-----  
**FIRMA**

## Anexo 4.- Certificado



### CERTIFICACIÓN

Latacunga, 30 de septiembre del 2018

Yo, **Ángel Salvador Hernández Moreno**, con cédula de identidad N° 1757109366 e Investigador Principal del Proyecto de Investigación Generativa "Evaluación del proceso de secado en Horno Microondas de Semillas de Interés Agrícola de la Provincia de Cotopaxi" autorizo a la MSc. Silvia Jeaneth Bravo Mullo (Cl. 0502437122) para que emplee los datos experimentales obtenidos en dicho proyecto para el desarrollo de un "Modelo basado en Machine Learning para la predicción de variables empleadas en el secado de semillas en horno microondas". Dicho modelo será desarrollado por las estudiantes de décimo ciclo de la Carrera de Ingeniería en Informática y Sistemas Computacionales Silvia Carolina Cárdenas Quinapaxi (Cl. 1723170658) y Vanessa Estefanía Vela Pérez (Cl. 1804664496) bajo la Dirección de la MSc. Silvia Jeaneth Bravo Mullo.

Y para que así conste para los fines pertinentes, firmo la presente en Latacunga, a 30 de Septiembre de 2018.

PhD. Ángel Salvador Hernández Moreno

Investigador Principal del Proyecto

## Anexo 5.- Código de la variable tiempo de secado aplicando Redes Neuronales

```

import pandas as pd
import numpy as np
from sklearn.neural_network import MLPRegressor
from sklearn import linear_model
from sklearn.metrics import r2_score

datos=pd.read_csv("data2.csv")
x=datos["temperatura"]
y=datos["tiempo_secado"]

X=x[:, np.newaxis]
while True:

    from sklearn.model_selection import train_test_split
    X_train, X_test, y_train, y_test = train_test_split(X,y, test_size=0.20, train_size=0.80)

    mlr=MLPRegressor(solver='lbfgs', alpha=1e-5,hidden_layer_sizes=(5,5), random_state=1)

    mlr.fit(X_train, y_train)

    print ("score" , mlr.score(X_train, y_train))
    if mlr.score (X_train, y_train)>0.9954:
        break

#Resultados de las predicciones.
print ("-----Predicción de la variable: Tiempo de secado-----")
print ("*Temperatura de 35 °C", mlr.predict(35))
print ("*Temperatura de 45 °C", mlr.predict(45))
print ("*Temperatura de 55 °C", mlr.predict(55))
y_pred=mlr.predict(X_test)

#precision del modelo
lr = linear_model.LinearRegression()
lr.fit(X_train, y_train)
y_pred = lr.predict(X_test)
print("Precisión del modelo:")
print(lr.score(X_train, y_train))

# Puntaje de Varianza. El mejor puntaje es un 1.0
regr = linear_model.LinearRegression()
regr.fit(X_train, y_train)
y_pred = regr.predict(X_train)
print("-----Puntaje de varianza-----")
print('El puntaje de varianza es: %.2f' % r2_score(y_train, y_pred))

```

## Anexo 6.- Código de la variable consumo de energía aplicando Redes Neuronales

```

import pandas as pd
import numpy as np
from sklearn.neural_network import MLPRegressor
from sklearn import linear_model
from sklearn.metrics import r2_score

datos=pd.read_csv("data2.csv")
x=datos["temperatura"]
y=datos["consumo_energia"]

X=x[:, np.newaxis]
while True:

    from sklearn.model_selection import train_test_split
    X_train, X_test, y_train, y_test = train_test_split(X,y, test_size=0.20, train_size=0.80)

    mlr=MLPRegressor(solver='lbfgs', alpha=1e-5,hidden_layer_sizes=(5,5), random_state=1)

    mlr.fit(X_train, y_train)

    print ("score" , mlr.score(X_train, y_train))
    if mlr.score (X_train, y_train)>0.9954:
        break

#Resultados de las predicciones.
print ("-----Predicción de la variable: Tiempo de secado-----")
print ("*Temperatura de 35 °C", mlr.predict(35))
print ("*Temperatura de 45 °C", mlr.predict(45))
print ("*Temperatura de 55 °C", mlr.predict(55))
y_pred=mlr.predict(X_test)

#precision del modelo
lr = linear_model.LinearRegression()
lr.fit(X_train, y_train)
y_pred = lr.predict(X_test)
print("Precisión del modelo:")
print(lr.score(X_train, y_train))

# Puntaje de Varianza. El mejor puntaje es un 1.0
regr = linear_model.LinearRegression()
regr.fit(X_train, y_train)
y_pred = regr.predict(X_train)
print("-----Puntaje de varianza-----")
print('El puntaje de varianza es: %.2f' % r2_score(y_train, y_pred))

```

## Anexo 7.- Código de la variable tasa de germinación aplicando Redes Neuronales

```

import pandas as pd
import numpy as np
from sklearn.neural_network import MLPRegressor
from sklearn import linear_model
from sklearn.metrics import r2_score

datos=pd.read_csv("data2.csv")
x=datos["temperatura"]
y=datos["tasa_germinacion"]

X=x[:, np.newaxis]
while True:

    from sklearn.model_selection import train_test_split
    X_train, X_test, y_train, y_test = train_test_split(X,y, test_size=0.20, train_size=0.80)

    mlr=MLPRegressor(solver='lbfgs', alpha=1e-5,hidden_layer_sizes=(5,5), random_state=1)

    mlr.fit(X_train, y_train)

    print ("score" , mlr.score(X_train, y_train))
    if mlr.score (X_train, y_train)>0.9954:
        break

#Resultados de las predicciones.
print ("-----Predicción de la variable: Tiempo de secado-----")
print ("*Temperatura de 35 °C", mlr.predict(35))
print ("*Temperatura de 45 °C", mlr.predict(45))
print ("*Temperatura de 55 °C", mlr.predict(55))
y_pred=mlr.predict(X_test)

#precision del modelo
lr = linear_model.LinearRegression()
lr.fit(X_train, y_train)
y_pred = lr.predict(X_test)
print("Precisión del modelo:")
print(lr.score(X_train, y_train))

# Puntaje de Varianza. El mejor puntaje es un 1.0
regr = linear_model.LinearRegression()
regr.fit(X_train, y_train)
y_pred = regr.predict(X_train)
print("-----Puntaje de varianza-----")
print('El puntaje de varianza es: %.2f' % r2_score(y_train, y_pred))

```

## Anexo 8.- Código de la variable tiempo de secado aplicando Random Forests

```

import pandas as pd
import numpy as np
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import cross_val_score
from sklearn import linear_model
from sklearn.metrics import r2_score

datos=pd.read_csv("data2.csv")
x=datos["temperatura"]
y=datos["tiempo_secado"]
X=x[:, np.newaxis]

while True:
    from sklearn.model_selection import train_test_split
    X_train, X_test, y_train, y_test = train_test_split(X,y, test_size=0.20, train_size=0.80)

    mlr=RandomForestRegressor(n_estimators=10, random_state=42)

    mlr.fit(X_train, y_train)
    cv_results= cross_val_score(mlr,X_train, y_train)

    print (mlr.score(X_train, y_train))

    if mlr.score (X_train, y_train)>0.9954:
        break
print ("Predicción de la variable: Tiempo de secado")
print ("Temperatura de 35 °C", mlr.predict(35))
print ("Temperatura de 45 °C", mlr.predict(45))
print ("Temperatura de 55 °C", mlr.predict(55))
y_pred=mlr.predict(X_test)

msg="Random: %f (%f)" % (cv_results.mean(), cv_results.std())

#precision del modelo
lr = linear_model.LinearRegression()
lr.fit(X_train, y_train)
y_pred = lr.predict(X_test)
print("Precisión del modelo:")
print(lr.score(X_train, y_train))

# Puntaje de Varianza. El mejor puntaje es un 1.0
regr = linear_model.LinearRegression()
regr.fit(X_train, y_train)
y_pred = regr.predict(X_train)

```

## Anexo 9.- Código de la variable consumo de energía aplicando Random Forests

```

import pandas as pd
import numpy as np
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import cross_val_score
from sklearn import linear_model
from sklearn.metrics import r2_score

datos=pd.read_csv("data2.csv")
x=datos["temperatura"]
y=datos["consumo_energia"]
X=x[:, np.newaxis]

while True:
    from sklearn.model_selection import train_test_split
    X_train, X_test, y_train, y_test = train_test_split(X,y, test_size=0.20, train_size=0.80)

    mlr=RandomForestRegressor(n_estimators=10, random_state=42)

    mlr.fit(X_train, y_train)
    cv_results= cross_val_score(mlr,X_train, y_train)

    print (mlr.score(X_train, y_train))

    if mlr.score (X_train, y_train)>0.9954:
        break
print ("Predicción de la variable: Tiempo de secado")
print ("Temperatura de 35 °C", mlr.predict(35))
print ("Temperatura de 45 °C", mlr.predict(45))
print ("Temperatura de 55 °C", mlr.predict(55))
y_pred=mlr.predict(X_test)

msg="Random: %f (%f)" % (cv_results.mean(), cv_results.std())

#precision del modelo
lr = linear_model.LinearRegression()
lr.fit(X_train, y_train)
y_pred = lr.predict(X_test)
print("Precisión del modelo:")
print(lr.score(X_train, y_train))

# Puntaje de Varianza. El mejor puntaje es un 1.0
regr = linear_model.LinearRegression()
regr.fit(X_train, y_train)
y_pred = regr.predict(X_train)

```

## Anexo 10.- Código de la variable tasa de germinación aplicando Random Forests

```

import pandas as pd
import numpy as np
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import cross_val_score
from sklearn import linear_model
from sklearn.metrics import r2_score

datos=pd.read_csv("data2.csv")
x=datos["temperatura"]
y=datos["tasa_germinacion"]
X=x[:, np.newaxis]

while True:
    from sklearn.model_selection import train_test_split
    X_train, X_test, y_train, y_test = train_test_split(X,y, test_size=0.20, train_size=0.80)

    mlr=RandomForestRegressor(n_estimators=10, random_state=42)

    mlr.fit(X_train, y_train)
    cv_results= cross_val_score(mlr,X_train, y_train)

    print (mlr.score(X_train, y_train))

    if mlr.score (X_train, y_train)>0.9954:
        break
print ("Predicción de la variable: Tiempo de secado")
print ("Temperatura de 35 °C", mlr.predict(35))
print ("Temperatura de 45 °C", mlr.predict(45))
print ("Temperatura de 55 °C", mlr.predict(55))
y_pred=mlr.predict(X_test)

msg="Random: %f (%f)" % (cv_results.mean(), cv_results.std())

#precision del modelo
lr = linear_model.LinearRegression()
lr.fit(X_train, y_train)
y_pred = lr.predict(X_test)
print("Precisión del modelo:")
print(lr.score(X_train, y_train))

# Puntaje de Varianza. El mejor puntaje es un 1.0
regr = linear_model.LinearRegression()
regr.fit(X_train, y_train)
y_pred = regr.predict(X_train)

```