

**DISEÑO DE UN SISTEMA DE RIEGO PARA
GRANJA INTELIGENTE BASADO EN
INTERNET DE LAS COSAS**

DISEÑO DE UN SISTEMA DE RIEGO PARA GRANJA INTELIGENTE BASADO EN INTERNET DE LAS COSAS



**UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS**

PROPONENTE
Ing. Gabriel Andrés Alzate Acuña

DIRECTOR
PhD. Roberto Ferro Escobar

Maestría en Ciencias de la Información y las Comunicaciones,
Universidad Distrital Francisco José de Caldas, Bogotá
2020.

El momento en el que sientas que a lo mejor caminas por la calle desnudo, llevando al aire demasiado de tu corazón y de tu mente y de lo que hay en tu interior, enseñando demasiado de ti mismo. Ese es el momento en el que tal vez estés empezando a hacerlo bien.

Neil Gaiman

ÍNDICE GENERAL

Índice de figuras	IX
Índice de cuadros	XI
Glosario	XIII
1 Introducción	1
1.1 Antecedentes	1
1.2 Planteamiento del problema de Investigación	3
1.3 Pregunta de Investigación.	3
1.4 Objetivos	4
1.4.1 Objetivo General.	4
1.4.2 Objetivos Específicos	4
1.5 Justificación.	5
1.6 Alcances y Limitaciones.	6
1.6.1 Alcances	6
1.6.2 Limitaciones.	6
2 Marco Teórico	7
2.1 Introducción	7
2.2 Conceptos Teóricos	9
2.2.1 Conceptos sobre Riego y modelos del mismo	9
2.2.2 Requerimientos de agua de los cultivos	10
2.2.3 Factores que afectan la ET	11
2.2.4 Métodos para determinar la evapotranspiración.	13
2.2.5 Ecuación de FAO Penman-Monteith	15
2.2.6 Ecuación de Hargreaves para el cálculo de la evapotranspiración	16
2.3 Gestión del conocimiento.	16
2.3.1 Modelos de gestión del conocimiento	17
2.3.2 Desarrollo del proceso de gestión del conocimiento	20
2.3.3 Etapas que componen el proceso de gestión del conocimiento	21
2.4 Inteligencia Computacional.	22
2.4.1 Tipos de aprendizaje.	23
2.4.2 Herramientas de Inteligencia Computacional	23
2.5 El paradigma IoT	27
2.5.1 Protocolos de comunicación IoT.	29
2.5.2 Computación en la niebla (Fog Computing)	33
2.5.3 Computación en el techo (Roof Computing)	34
2.5.4 Criterios para realizar la implementación de un sistema IoT.	36
2.5.5 Plataformas IoT	38

2.5.6	Estándar W3C para SSN	40
2.6	Estudios similares.	44
2.6.1	Sensado de parámetros relevantes en agricultura	44
2.6.2	Propuestas de arquitectura	45
2.6.3	Estimación de riego	48
3	Datos Y Métodos	51
3.1	Introducción	51
3.2	Diseño de la investigación	51
3.3	VARIABLES	52
3.4	Población y muestra	52
3.5	Instrumentos y procedimientos.	52
3.6	Análisis de los datos.	53
3.6.1	Sensores	53
3.6.2	Captura de la información de los sensores	55
3.6.3	Calculo de la evapotranspiración	56
4	Desarrollo	61
4.1	Fuentes de conocimiento	63
4.1.1	Estación meteorológica	64
4.2	Backend - Estructura general	67
4.2.1	Generalidades	67
4.2.2	Estructura de datos	68
4.2.3	Base de datos	69
4.2.4	Configuración del servidor.	71
4.3	Inteligencia Artificial - Cálculo de la evapotranspiración	73
4.3.1	Cálculo por medio de la ecuación de Hargreaves.	73
4.3.2	Desarrollo del sistema de inteligencia computacional de estimación de riego	74
4.3.3	Análisis del modelo desarrollado LGB on Elastic-Net	80
4.3.4	Evaluación de la estabilidad del modelo	83
4.4	Frontend - Dashboard	85
4.4.1	index.html	85
4.4.2	Imágenes del sistema	85
4.5	Documento técnico de uso del Sistema de Riego	86
4.5.1	Fuentes de conocimiento	88
4.5.2	Identificación de variables	88
4.5.3	Selección de variables relevantes.	89
4.5.4	Almacenamiento estructurado.	89
4.5.5	Combinación y transferencia	90
4.5.6	Asimilación y uso	90
5	Conclusiones	91
6	Trabajos Futuros	93
6.0.1	Trabajos en ingeniería electrónica	93
6.0.2	Trabajos en ingeniería de sistemas	93

6.0.3	Trabajos en ingeniería industrial	94
6.0.4	Tesis de maestría	94
	Bibliografía	97
A	Anexo: Producción Bibliográfica	101
A.0.1	Titulo	101
A.0.2	Resumen	101
A.0.3	Palabras clave	102
A.0.4	Detalles	102

ÍNDICE DE FIGURAS

2.1 Diagrama de tres fases presentes en el suelo	10
2.2 Representación de la estoma	11
2.3 Evaporación y transpiración durante crecimiento de un cultivo	12
2.4 Factores que afectan la ET con referencia a conceptos relacionados	13
2.5 Representación de ET_0 , ET_c y ET_{caj}	14
2.6 Pirámide del conocimiento.	18
2.7 Arquitectura del modelo integrado-situacional.	19
2.8 Espiral de la gestión del conocimiento	20
2.9 Creación y transformación de conocimiento	21
2.10 Proceso de gestión del conocimiento	22
2.11 Representación digital de una neurona	24
2.12 Proceso de clasificación con SMV	26
2.13 Representación gráfica de un conjunto difuso	27
2.14 Proyectos en desarrollo en IoT	29
2.15 Empresas en agricultura inteligente	30
2.16 Esquema protocolo MQTT	33
2.17 Representación de la niebla en Fog Computing	34
2.18 Arquitectura computación en la niebla	35
2.19 Ubicación de Roof Computing	36
2.20 Cuatro funcionalidades Roof Computing	36
2.21 Secuencia de implementación de un proyecto IoT	37
2.22 Tecnologías y casos de uso	42
2.23 Módulos de la Ontología SSN	43
2.24 Vista de la Ontología SSN en clases y propiedades	44
2.25 Arquitectura de manejo de granja bajo el paradigma IoT	47
2.26 Arquitectura de SmartFarmNet	48
2.27 Sistema de calculo para necesidades de riego	50
4.1 Arquitectura propuesta para riego inteligente	61
4.2 Proceso de gestión del conocimiento	63
4.3 Datos meteorológicos y asociados	64
4.4 Entrada / Salida de información del HOST	66
4.5 Backend implementado	67
4.6 Nombres y propiedades relevantes según <i>clima y pronostico</i>	68
4.7 Estandarización de nombres de variables según ontología W3C	69
4.8 Base de datos implementada para almacenamiento, presentación y estimación de riego	70
4.9 Algoritmo de captura, limpieza y almacenamiento de datos	71

4.10 Matriz de asociación	75
4.11 Importancia de Permutación normalizada	78
4.12 Flujo de trabajo del modelo Light Gradient Boosting	81
4.13 Tabla de elevación del modelo desarrollado	84
4.14 Banner del sistema	85
4.15 Encabezado del sistema	86
4.16 Variables del sistema	86
4.17 Variables del sistema	87
4.18 Pie de página del sistema	87
4.19 Proceso de gestión del conocimiento	88

ÍNDICE DE CUADROS

2.1 Radiación solar según latitud. Fuente: Propia adaptada de (Hargreaves and Allen, 2003)	17
2.2 Comparativa de plataformas IoT de código abierto	41
3.1 Sensores certificados por la empresa Davis en EEUU	54
3.2 Frecuencia de datos de sensores	55
3.3 Sensores de temperatura y humedad	56
3.4 Sensores radiación solar	57
3.5 Sensor de velocidad del viento	58
3.6 Sensor de velocidad del viento	59
3.7 Datos provenientes de sensores. Fuente: Propia.	60
4.1 Carga estadística de las variables. Fuente propia	77
4.2 Resultados de la prueba "Importancia de permutación" totales y normalizados. Fuente propia	78
4.3 Diferentes modelos entrenados y considerados	80
4.4 Diferentes modelos entrenados y considerados parte 2	80
4.5 Diferentes modelos entrenados y considerados parte 3	81
4.6 Kc para diferentes tipos de cultivo en diferentes etapas de crecimiento. Fuente: FAO (FAO, 2017)	90

GLOSARIO

ET_0 Evapotranspiración de referencia.

ET_c Evapotranspiración de cultivo estandar.

ET_{caj} Evapotranspiración de cultivo bajo condiciones no estándar.

K_c Coeficiente de cultivo.

K_s Coeficiente de estrés hídrico.

ET Evapotranspiración.

FAO Organización de las Naciones Unidas para la Alimentación y la Agricultura.

ICID International Commission on Irrigation and Drainage.

IEEE Institute of Electrical and Electronics Engineers.

IoT Internet de las Cosas.

OMM Organización Meteorológica Mundial.

OWL Web Ontology Language.

SIAC Sistema de Información Ambiental de Colombia.

SSN Semantic Sensor Network.

W3C W3C Semantic Sensor Network Incubator Group.

WSN Red de Sensores Inalámbricos (Wireless Sensor Network).

1

INTRODUCCIÓN

1.1. ANTECEDENTES

El desarrollo económico y social de un país depende de la posibilidad de abastecerse de alimentos estableciendo infraestructuras de producción en su territorio o formando redes de intercambios comerciales con otros países, además se considera un motor de desarrollo la producción de excedentes de alimentos para destinarlos a exportación.

El aumento de la población asistida por su derecho de alimentarse adecuadamente ha hecho que los recursos para la producción agrícola como la tierra y el agua sean cada vez menos empujando el límite de la frontera agrícola, lo que significa una disminución en la cantidad de selvas y ecosistemas nativos de las regiones sustituyéndolas por cultivos, práctica que va en contravía del desarrollo sustentable ([de las Naciones Unidas para la Alimentación y la Agricultura, 2015](#)). Esta disminución de recursos ha generado procesos de tecnificación de la producción de alimentos en áreas como la genética con los transgénicos, las granjas urbanas como alternativa de usos del suelo o la inclusión de nuevas tecnologías en la cadena de valor de cada producto todo esto en búsqueda de aumentar la cantidad de alimentos producidos; esto no deja de ser una práctica inacabada pues las demandas de alimentos siempre van a ser inherentes al ser humano.

La aplicación de agua a suelos agrícolas con el propósito de regar cultivos es una de estas tecnologías y la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO) ha estimado que un manejo adecuado de los recursos de agua representará un aumento más importante en los rendimientos de los cultivos agrícolas que **el conjunto** de todas las otras prácticas de producción juntas ([FAO, 2017](#)). En Colombia en el 90% de los cultivos el riego no se encuentra tecnificado o lo está parcialmente hecho por medio de tecnologías rudimentarias, con poca o nula incorporación de sistemas robustos especializados pues la tecnología de riego debe responder a la combinación específica del suelo, el cultivo y el clima, entre otras variables. El no desarrollarlo de forma adecuada genera problemas como pérdidas de agua, es decir, una baja eficiencia en el aprovechamiento del recurso, lavado de nutrientes naturales bajo la zona donde se desarrollan las raíces, bajos rendimientos en los cultivos por falta o exceso de agua en diferentes partes del cultivo bajando el rendimiento de estos ([Gurovich, 2012](#)).

La transferencia tecnológica es el proceso por el cual al surgir de nuevas tecnologías y procesos su aplicación y desarrollo se traslada a diferentes sectores con el fin de robustecer las actividades propias del mismo. Entre esas tecnologías aparece el Internet de las Cosas (IoT) como paradigma que vincula el levantamiento de datos, su transporte y posterior aplicación para generar información y conocimiento (Zanella et al., 2014).

El IoT es un paradigma de desarrollo e implementación mundial y a generado un movimiento la nivel local en Colombia, destacan el Meetup *IoT Colombia*, creado por una agrupación llamada HackBo, empresas como Zebra Technologies, Solutek, Ubidots, AZlogica, AR-Smart o la empresa francesa Sigfox con presencia Colombia con su plataforma de red para IoT.

Con el interés de formar la triple hélice (que es el conjunto formado por la academia, industria y gobierno) el gobierno a través del Ministerio de las Tecnologías de la Información y Comunicaciones (MINTIC) apoyo financiera y logísticamente la unión de las universidades Javeriana, Autónoma de Bucaramanga, Tecnológica de Bolívar y Santo Tomas junto a las empresas Lógica SAS, Banco Pichincha, Hospital San Ignacio, Totto y Zona Franca de Bogotá para formar el Centro de Excelencia y Apropiación de Internet de las Cosas (CEA-IoT) con el fin de generar desarrollo técnico y tecnológico basado en IoT para la ciudad y el país. Esto sobre plataformas tecnológicas de sus aliados tecnológicos Microsoft, Intel y Hewlett Packard (de la Información y las Comunicaciones, 2017), este es el paso mas ambicioso dado por el gobierno para promover el uso y desarrollo de IoT.

Mientras las soluciones basadas en IoT brindan un ecosistema rico en aplicaciones y entendimiento, vienen de la mano de una nueva variedad de complejidad y confusión. En su reporte de desarrollo de IoT, IoT Analytics presenta 5 etapas en el desarrollo de una aplicación específica, que son el dispositivo, la comunicación, los servicios en la nube, las aplicaciones y la seguridad (Analytics, 2016).

Con base en lo anterior toma relevancia el uso de nuevas tecnologías para darle mayores rendimientos a la producción agrícola de la mano del IoT, tomando información ambiental y climatológica tanto de forma local como general para estimar las necesidades hídricas de los cultivos. Este postulado es la base del presente trabajo, en donde se hace una revisión de las necesidades particulares del riego de los cultivos, la tecnología con la que se realiza actualmente y los beneficios que presenta para luego proponer un sistema basado en el paradigma de IoT y de forma conjunta con la gestión del conocimiento y la inteligencia computacional crear una herramienta nueva que permita generar mayor rendimiento en los cultivos por medio del control del uso del agua en él.

1.2. PLANTEAMIENTO DEL PROBLEMA DE INVESTIGACIÓN

El aumento de la población a hecho crecer la demanda de alimentos motivando que los recursos como el agua o el suelo se vuelvan cada vez mas escasos y ha llevado al abuso de los mismos en varias zonas geográficas del mundo; la necesidad de manejar de una forma racional estos recursos por medio de un análisis de los mismos es una de las tareas pendientes no solo en Colombia sino a nivel mundial.

Para lograr un rendimiento productivo acorde a las necesidades de la nación es necesario controlar el desarrollo de la frontera agrícola, proceso apoyado en la creación de redes de riego y mejorando las existentes, con el fin de asegurar un uso eficiente de los recursos con que cuentan los agricultores.

La falta de transferencia tecnológica hacia el sector rural en el país ha sido una constante el los últimos 50 años (Gurovich, 2012), a pesar de los esfuerzos realizados por la Corporación Colombiana de Investigación Agropecuaria (CORPOICA) en realizar labor de extensión y transferencia tecnológica por lo que los esfuerzos encaminados a mejorar la productividad y rendimiento agrícola son recibidos con una amplia aceptación por parte de la comunidad (Mora and Arteaga, 2014). La Organización de las Naciones Unidas para la Alimentación y Agricultura (FAO, 2017) ha estimado que un manejo adecuado del agua entre riego y drenaje generará un rendimiento mayor en los cultivos que el conjunto de todas las otras prácticas de producción, además los procesos agrícolas consumen la mayor cantidad de este líquido en el mundo, además la demanda hídrica de la atmósfera es casi continua mientras que la lluvia y el riego ocurren ocasionalmente y en forma irregular. Para sobrevivir en períodos secos los cultivos debe estar capacitados para hacer uso del agua contenida en el espacio poroso del suelo.

1.3. PREGUNTA DE INVESTIGACIÓN

Se genera las siguiente pregunta de investigación que define el desarrollo del proyecto:

¿Qué debe tener un desarrollo tecnológico basado en IoT para ser apropiado por agricultores y mejorar el sistema de riego con el que manejan sus cosechas?

Esta pregunta de investigación genera una hipótesis.

- El correcto modelo de riego inteligente basado en Internet de las Cosas debe hacer uso del conocimiento generado previamente y utilizar la inteligencia computacional para estimar el riego de un cultivo y así controlar el mismo.

1.4. OBJETIVOS

Para darle solución al problema presentado, responder la pregunta de investigación y verificar la hipótesis presentada se proponen los siguientes objetivos del proyecto.

1.4.1. OBJETIVO GENERAL

Diseñar un sistema de riego inteligente aplicado a procesos agrícolas que integre variables analizadas de forma computacional provenientes de sensores de luz, humedad, temperatura y condición del suelo mediante el paradigma de IoT y la gestión del conocimiento.

1.4.2. OBJETIVOS ESPECÍFICOS

- Investigar la forma de modelar computacionalmente algunas variables asociadas al riego.
- Diseñar un sistema inteligente IoT que integre la información proveniente de sensores que registren las variables provenientes de una granja tradicional.
- Modelar un servicio web para el riego de granjas inteligentes basado en internet de las cosas que genere información valiosa por medio de inteligencia computacional.
- Crear un documento técnico que use la gestión del conocimiento para integrar las diferentes variables y su interrelación permita mejorar la forma en que se controla el riego en las granjas inteligentes en Colombia.

1.5. JUSTIFICACIÓN

La FAO es la Organización de las Naciones Unidas para la Alimentación y la Agricultura, entidad ejecutora de proyectos a nivel mundial nacida por decisión de la mayoría de naciones del mundo para solucionar los grandes inconvenientes derivados del proceso agrícola. Al ser parte de las Naciones Unidas comparte su visión, entre ellos los 17 objetivos de desarrollo sostenible, un esfuerzo en dar luces sobre el destino de los esfuerzos de crecimiento humano. Estos objetivos cubren las áreas que se pretenden priorizar para eliminar la desigualdad de condiciones de vida entre las naciones del mundo y busca al adherirse a estos lineamientos un mundo más justo, prospero, pacífico y sostenible.

La batalla para acabar con el hambre y la pobreza debe centrar sus esfuerzos en las zonas rurales, pues es allí donde vive el 80% de las personas que padecen hambre y los pobres del mundo, además la agricultura es el mayor empleador del mundo y el crecimiento de las economías agrarias de bajos ingresos es dos veces más eficaz que el crecimiento en otros sectores. Según la FAO es posible erradicar el hambre para 2030 con un esfuerzo a favor de los pobres en agricultura sostenible, desarrollo rural y medidas de protección social, para sacar a las personas de la subalimentación crónica y la pobreza ([de las Naciones Unidas para la Alimentación y la Agricultura, 2015](#)). Para que el sector agrícola colombiano sea parte en la implementación de objetivos de desarrollo sostenible es necesario dar un salto tecnológico a nivel rural en la cadena de producción agrícola mediada por las tecnologías de la información y las comunicaciones, que apalanque al agricultor en la mejora de producción y procesos que estén vinculados al comercio. Esto es necesario para el productor, comerciante y consumidor de productos agrícolas, lo que abarca básicamente todos los renglones de la sociedad.

Se evidencia la necesidad del desarrollo e implementación de tecnologías para que el sector agrícola colombiano logre el nivel de desarrollo y los estándares de calidad que requiere el comercio tanto a nivel nacional como mundial como medio para mejorar las condiciones de vida del agricultor y brindar al país un método de lograr la soberanía alimentaria.

En Colombia falta una recolección y reordenamiento de los componentes básicos que modelan de forma racional el riego y el drenaje: datos hidrológicos y agro-climáticos confiables, características hidromecánicas de los suelos ([Mora and Arteaga, 2014](#)). El resultado económico proyectado de los proyectos implementados en Colombia no ha sido satisfactorio, incluso décadas después de la construcción de embalses debido a la no implementación a nivel micro de proyectos que usen de forma útil el agua embalsada y distribuida, además los proyectos realizados se han hecho con una serie de supuestos técnicos en donde si un solo parámetro de estos no es adecuado, la implementación y ejecución no resulta ser una buena inversión, esto significa que no es posible tener un sistema de riego estándar para todos los cultivos sino que cada uno debe tener consideraciones especiales.

A diferencia con otras tecnologías las inversiones en riego son escalonadas y progresivas en el tiempo, cualquier inversión realizada siempre llevara un aumento de rendimiento ([Gurovich, 2012](#)), si otro factor de la producción no pasa a ser limitante.

1.6. ALCANCES Y LIMITACIONES

1.6.1. ALCANCES

Este trabajo pretende diseñar sistema de riego inteligente aplicado a procesos agrícolas que integre variables analizadas de forma computacional provenientes de sensores de luz, humedad, temperatura y condición del suelo donde la recolección de datos se realiza por una red de sensores inalámbricos.

Específicamente, se pretende definir los parámetros y criterios en un sistema de riego inteligente y la captura de los datos de entrada del modelo a través de sensores. Posteriormente se modelará un servicio web para el riego de una granja inteligente.

Se creará un documento técnico que use la gestión del conocimiento para integrar las diferentes variables tenidas en cuenta en el diseño.

1.6.2. LIMITACIONES

No se realizará la implementación de hardware de sensado, se dará uso de sensores comerciales desarrollando sus circuitos de acondicionamiento y adicionándoles formas de transmitir datos hacia un servicio web IoT.

La limitación económica la produce el apoyo institucional con el pago de plataformas tanto de software como de hardware, pues si los costos exceden el presupuesto asignado al mismo se vuelve inviable el uso de alguna tecnología, decantando soluciones basadas en software y hardware libre que hagan viable económicamente un piloto del proyecto.

La documentación técnica es una limitante, pues en Colombia no se han desarrollado análisis de tecnologías como IoT en cadenas de valor en el sector agropecuario, así que toda la información pertinente saldrá de experimentación dentro del proyecto con base en las anteriores limitaciones.

El hardware de sensado se asume realizado con componentes comerciales, pues existen líneas de sensores con resolución mayor pero solo se venden con fines militares.

2

MARCO TEÓRICO

2.1. INTRODUCCIÓN

El proyecto *Diseño de un sistema de riego para granja inteligente basado en Internet de las Cosas* abarca diferentes áreas del conocimiento a simple vista no interrelacionadas entre sí; al proponer un sistema que reconozca variables climáticas y ambientales se contextualizará al lector en los conceptos de riego y afines al mismo, dando una perspectiva del desarrollo actual de este proceso y al enmarcar el proyecto dentro de los sistemas de gestión del conocimiento se ampliará este en la medida que se necesite su desarrollo. El núcleo del proyecto es el sistema desarrollado bajo el paradigma de Internet de las Cosas, el cual es abordado en un inicio de manera integral y luego desde sus diferentes componentes para de esta forma cubrir los tópicos que dan alcance al proyecto.

El marco teórico propuesto inicia presentando los conceptos de riego y los modelos del mismo, en esta sección se define un parámetro de relevancia llamado *evapotranspiración* para luego explicar que es el suelo y sus características por ser el medio a través del cual se lleva agua a las raíces del cultivo y seguir con los requerimientos de agua de los cultivos, tópico que denota la forma y cantidad de las demandas hídricas de los cultivos.

Explicado esto se presenta a profundidad el concepto de evapotranspiración y los factores que la afectan y los métodos para determinarla con todas sus variaciones, este proceso será emulado en el proyecto con herramientas de inteligencia computacional. Entre estos métodos destaca la ecuación de FAO Penman-Monteith la cual es explicada y contextualizada para resaltar su relevancia.

Luego del contexto ambiental en el que se desenvolverán las tecnologías propuestas se precisa definir el contexto organizativo que se le darán a los datos reportados, para este propósito se define la gestión del conocimiento y los diferentes modelos que la abordan. Luego se presenta la forma de desarrollar este proceso y las fases o etapas que se siguen para lograrlo.

El paradigma IoT es el siguiente tópico en ser abordado partiendo de la base conceptual y presentando las diferentes definiciones que le han dado a lo largo del mundo

debido a que no hay un consenso de su definición, se abordan las aplicaciones que se pueden desplegar en diferentes locaciones y la forma en que se interconectan.

Conociendo la arquitectura de la solución IoT se abordan los protocolos de comunicaciones presentando sus fortalezas y debilidades, destaca el protocolo MQTT para la transmisión de datos, se abordan las formas en que pueden ser tratados estos desde la computación en la niebla hasta la computación de techo. La siguiente sección presenta los criterios con los cuales se realiza la implementación de un sistema IoT desde los diagramas de uso hasta la proyección final de socios interesados, se presentan las diferentes plataformas desarrolladas para el manejo de grandes volúmenes de datos IoT.

Para terminar se presentan los trabajos desarrollados a lo largo del mundo afines al proyecto y sus alcances, con esto se tiene un contexto amplio del problema y solución propuesta.

2.2. CONCEPTOS TEÓRICOS

2.2.1. CONCEPTOS SOBRE RIEGO Y MODELOS DEL MISMO

El término riego se refiere a un conjunto de métodos de aplicación del agua al suelo de forma que el agua se distribuya de forma homogénea. El diseño y proyección de las situaciones específicas del suelo, clima, disponibilidad de agua, distritos de riego y manejo de los actuales llevan características propias por cada caso evaluado, el conocimiento de estos procesos y la interacción de sus partes se establece en la producción agrícola, lo cual da información para predecir el comportamiento del cultivo y diseñar un sistema de riego que aumente el rendimiento y mejore la calidad de los productos finales al asegurar al componente vegetal una adecuada cantidad de agua.

Para aplicar agua a un cultivo el agricultor debe responder cuatro preguntas imprescindibles, estas son ¿Porqué regar? Cual es la motivación de rendimiento al controlar el riego. ¿Cuándo regar? Es decir, la frecuencia y criterio para elegir esa frecuencia. ¿Cuánto regar? En función de tiempo y volumen del caudal. ¿Cómo regar? Que responde la forma de aplicación del agua al suelo. Esto define el riego como la aplicación oportuna y uniforme de agua a un perfil del suelo para reponer en este el agua consumida por los cultivo entre dos riegos consecutivos (Gurovich, 2012), teniendo en cuenta que no se riega la superficie del suelo, sino que se esta regando el *perfil* en profundidad, ya que el objetivo es administrar agua al espacio donde se encuentran las raíces de la planta, un riego eficiente no humedece uniformemente la superficie del suelo sino que almacena agua homogéneamente en donde se encuentran las raíces del mismo.

La pérdida de agua en la planta se realiza por dos procesos diferentes, el uno libera agua a la atmósfera a través de la transpiración, el segundo desde la superficie del suelo a la atmósfera por el proceso de evaporación, los dos procesos no solo se relacionan con la planta sino por el clima, tipo de suelo, etapa de desarrollo del cultivo y se conoce conjuntamente a este proceso como *evapotranspiración*.

El riego superficial es la forma principal de aplicar agua a un cultivo para suplir en este el agua evapotranspirada desde el riego anterior cuyo objetivo principal del es mantener en forma constante y máxima la velocidad de transpiración durante toda la temporada de riego.

El riego se puede implementar por aspersión, sistema que consume mas energía que otros y tiene perdidas mayores de agua. Se puede implementar el riego mediante inundación de surcos por gravedad, tiene como desventaja que toca prepara en exceso el campo haciéndolo horizontal y al momento de implementarlo se genera un desperdicio de agua considerable. El sistema de irrigación por goteo es una tercera opción y se caracteriza por consumir menos agua y energía llevando el agua a donde se necesita específicamente. Este sistema de riego permite llevar fertilizantes a las plantas de forma directa.

EL SUELO

El suelo es un sistema de bastante complejidad constituido de materia en sus tres fases, sólido líquido y gaseoso. A saber material mineral y orgánico en la primer fase, agua en la segunda fase y la gaseosa se refiere al espacio poroso que hay entre partículas (Gurovich, 2012).

El término suelo corresponde a la capa mas superficial de la corteza terrestre formada principalmente por desintegración y descomposición de rocas y material animal ó vegetal a lo largo del tiempo. La composición del suelo vista desde las tres fases puede verse en la figura 2.1.



Figura 2.1: Diagrama de tres fases presentes en el suelo
Fuente:Propia

2.2.2. REQUERIMIENTOS DE AGUA DE LOS CULTIVOS

El consumo de agua de las plantas es desproporcionado con respecto al crecimiento vegetativo, esto significa que las plantas transfieren a la atmósfera la mayor parte (mas de un 90%) del agua que extraen del suelo ([de las Naciones Unidas para la Alimentación y la Agricultura, 2015](#)). El proceso de *transpiración* no es necesariamente vital de la planta y esta causada mas que por el crecimiento de los vegetales, por un gradiente de presión de vapor entre las hojas y la atmósfera que las rodea.

Al momento de determinar la frecuencia con que regar se pueden usar dos enfoques diferentes, uno desde el punto de vista del agua en el suelo y otra desde el consumo de agua de la planta incluyendo el proceso de transpiración de la planta y la evaporación del agua en el suelo, proceso descrito previamente como Evapotranspiración (ET).

La planta no tiene un comportamiento pasivo ya que tiene la propiedad de limitar su transpiración cuando es sometida a bajos niveles de agua y cerrar los estomas de sus hojas. Sin embargo, cuando limita este proceso se reduce al mismo tiempo el crecimiento potencial ya que los estomas que transpiran el vapor a la atmósfera sirven también para absorber anhídrico carbónico, que es el sustrato principal de la fotosíntesis y una transpiración menor genera un alza en la temperatura interna de la planta pudiendo producir un colapso celular. En conclusión lo que se busca es la regulación de la transpiración a través de los estomas de la planta, la representación esquemática del estoma se aprecia en la figura 2.2.

Un déficit hídrico en las hojas es una disminución de la velocidad de transpiración y

Atmósfera

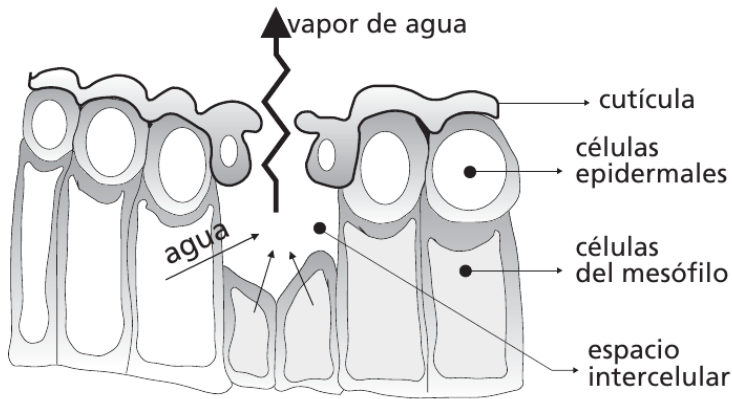


Figura 2.2: Representación de la estoma
Fuente:(FAO, 2017)

un aumento de la temperatura de la hoja. El cierre de las estomas se vincula con un bajo contenido de humedad en el suelo y con un alto grado de ET. Además un exceso de agua ocupara mucho volumen en el espacio poroso del suelo, esto debido a que el oxígeno molecular difícilmente se disuelve en agua impidiendo su acceso a la planta.

La evaporación y transpiración en la planta presente en el cultivo tiene ciclos que se basan en el tiempo de desarrollo de esta, siendo así se hace pertinente el análisis bajo parámetros propios de la planta, para ilustrar este ciclo, en la figura 2.3 se aprecian la repartición de la evapotranspiración en evaporación y transpiración durante el periodo de crecimiento de un cultivo anual.

Se han desarrollado varios métodos para medir la cantidad de agua presente en el suelo y se agrupan en dos tipos, los métodos directos y lo indirectos, en los primeros se separa directamente el agua del suelo en los segundos se aprovecha alguna propiedad física del sistema suelo-agua que varíe de acuerdo al contenido de agua.

El método directo mas extendido es el muestreo y posterior secado en estufa a 105 °C tomando muestras que dependen de la variabilidad del terreno y a diferentes profundidades del perfil donde se quiera llegar con el riego. Los métodos indirectos se basan en su gran mayoría en derivaciones de la conductividad eléctrica presente en materiales porosos a través de bloques de resistencia eléctrica o "bloques de Bouyoucos".

Con base en la información presentada se precisa que calcular los requerimientos de agua de un cultivo es sinónimo de determinar la evapotranspiración del mismo.

2.2.3. FACTORES QUE AFECTAN LA ET

Las variables principales a tener en cuenta al momento de calcular la ET se distinguen en tres grupos, las variables climáticas, los factores de cultivo y el manejo y condiciones ambientales.

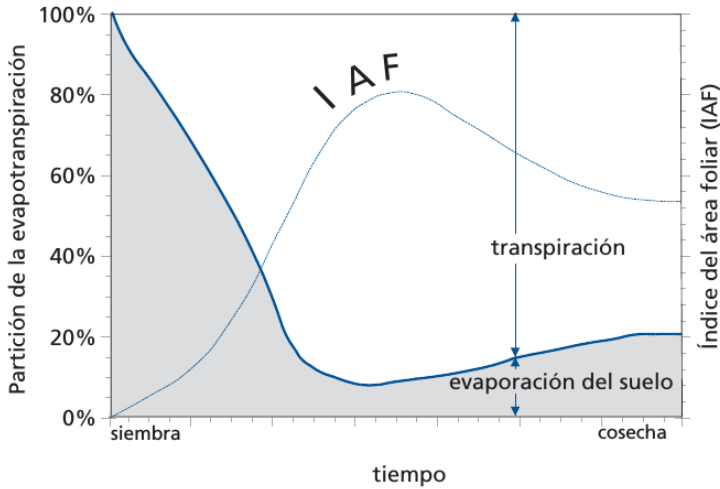


Figura 2.3: Evaporación y transpiración durante crecimiento de un cultivo

Fuente:FAO (FAO, 2017)

La ET se puede dividir en Evapotranspiración de referencia (ET_0), Evapotranspiración de cultivo estándar (ET_c) y Evapotranspiración de cultivo bajo condiciones no estándar (ET_{caj}). ET_0 es un término que relaciona el poder evaporante de la atmósfera, ET_c se refiere a la evapotranspiración en condiciones óptimas con la máxima producción y ET_{caj} es la evapotranspiración de un cultivo que crece bajo condiciones no óptimas como son presencia de plagas y enfermedades, salinidad del suelo baja fertilidad o exceso de agua. La relación de estos factores se presenta en la figura 2.4.

La ET_0 ocurre sin restricciones de agua en una superficie de referencia que corresponde a un cultivo hipotético de pasto con características específicas, los únicos factores que afectan ET_0 son los parámetros climáticos, por lo tanto, ET_0 es también un parámetro climático que puede ser calculado a partir de datos meteorológicos y expresa el poder evaporante de la atmósfera en una localidad y época del año específicas sin considerar las características del cultivo ni los factores del suelo.

La ET_c se refiere a la evapotranspiración de cualquier cultivo cuando se encuentra exento de enfermedades, con buena fertilización y que se desarrolla en parcelas amplias bajo óptimas condiciones de suelo y agua, y alcanza la máxima producción de acuerdo a las condiciones climáticas reinantes. La relación $\frac{ET_c}{ET_0}$ puede ser determinada experimentalmente para diferentes cultivos y es conocida como Coeficiente del Cultivo (K_c), se utiliza para relacionar ET_c a ET_0 de manera que $ET_c = K_c * ET_0$.

La ET_{caj} se refiere a la evapotranspiración de cultivos que crecen bajo condiciones ambientales y de manejo diferentes de las condiciones estándar. La evapotranspiración real del cultivo puede desviarse de ET_c debido a condiciones no óptimas como son la presencia de plagas y enfermedades, salinidad del suelo, baja fertilidad del suelo y limitación o exceso de agua. La ET_{caj} se calcula utilizando un coeficiente de estrés hídrico K_s o ajustando K_c a todos los otros tipos de condiciones de estrés y limitaciones ambien-

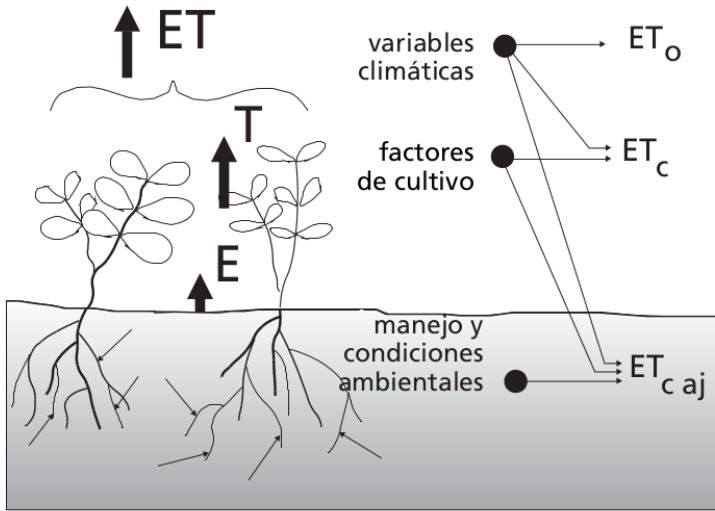


Figura 2.4: Factores que afectan la ET con referencia a conceptos relacionados
Fuente:FAO (FAO, 2017)

tales en la evapotranspiración del cultivo. Esta relación se aprecia en la figura 2.5.

2.2.4. MÉTODOS PARA DETERMINAR LA EVAPOTRANSPIRACIÓN

En mayo de 1990 la FAO organizó una consulta de expertos con la colaboración de la International Commission on Irrigation and Drainage (ICID) y la Organización Meteorológica Mundial (OMM) para revisar la metodología de estimación de las necesidades de agua en cultivos. El panel de expertos recomendó la adopción del método combinado de Penman-Monteith como el procedimiento estándar para medición de la evapotranspiración de referencia (FAO, 2017).

La ET se expresa normalmente en milímetros por unidad de tiempo. Esta unidad expresa la cantidad de agua perdida en una superficie cultivada en unidades de agua altura. La unidad de tiempo puede ser una hora, día, 10 días, mes o incluso un completo período de cultivo de un año.

Para determinar la ET experimentalmente se requieren instrumentos específicos y mediciones precisas de varios parámetros físicos. Generalmente los métodos aplicados en zona rural son caros y solo realizados apropiadamente por personal de investigación con un alto nivel de preparación. Entre estos métodos resaltan aquellos que buscan la *relación de energía entrante y saliente del suelo y la planta* por medio del principio de conservación de energía. Una aproximación diferente a este cálculo es realizado por el método de *balance de agua en el suelo*, que consiste en evaluar los flujos de agua que entran y salen de la zona radicular del cultivo dentro de un determinado periodo de tiempo (FAO, 2017).

Un método adicional es conocido como método por *lisímetros* donde se aísla la zona radicular que es la capa del suelo que alberga las raíces del cultivo y se controlan los

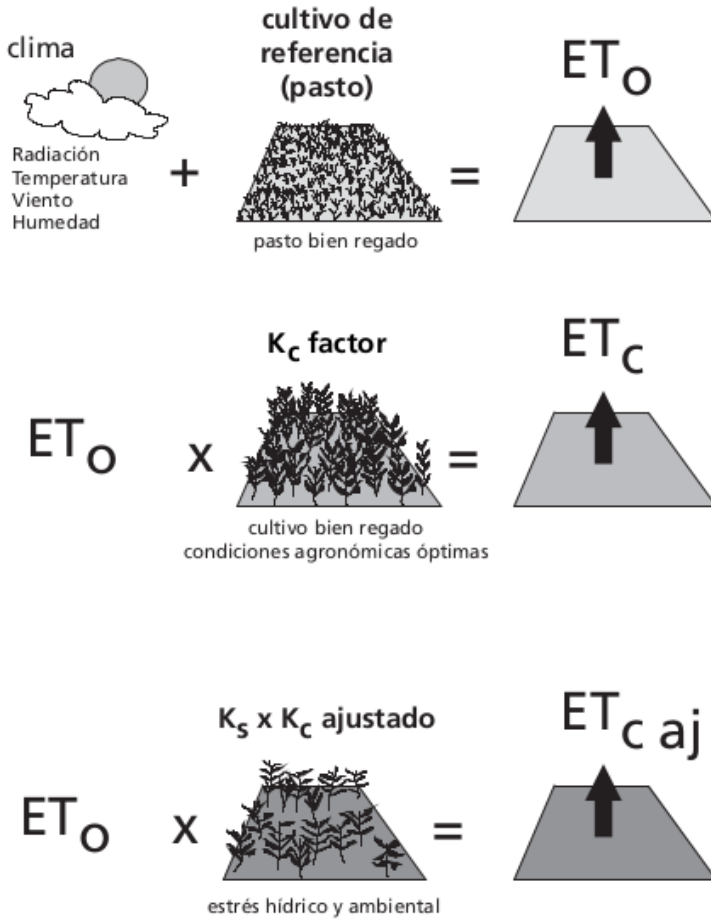


Figura 2.5: Representación de ET_0 , ET_c y $ET_{c\ aj}$
Fuente:FAO (FAO, 2017)

procesos que son difíciles de medir, los diversos términos en la ecuaciones del balance de agua en el suelo se pueden determinar con apreciable exactitud (Gurovich, 2012).

Debido a los inconvenientes para levantar información de campo fiable por los métodos descritos, la ET se calcula generalmente por medio de datos meteorológicos. Se han desarrollado varias ecuaciones empíricas y semi-empíricas. Como se ha mencionado por recomendaciones de la FAO se trabaja con el método FAO Penman-Monteith para el cálculo de ET_0 . La ET_c se determina luego usando el Coeficiente de cultivo (K_c), ya teniendo estos datos se procede a calcular la $ET_{c\ aj}$ ajustando la medida por medio del Coeficiente de estrés hídrico (K_s) o modificando el coeficiente de cultivo.

2.2.5. ECUACIÓN DE FAO PENMAN-MONTEITH

El método FAO Penman-Monteith fue desarrollado haciendo uso de la definición del cultivo de referencia como un cultivo hipotético con una altura asumida de 0,12 m, con una resistencia superficial de 70 s m^{-1} y un albedo de 0,23 (es el porcentaje de radiación que cualquier superficie refleja respecto a la radiación que incide sobre la misma) y que representa a la evapotranspiración de una superficie extensa de pasto verde de altura uniforme, creciendo activamente y adecuadamente regado, este se representa por la ecuación:

$$ET_0 = \frac{0,408\Delta(R_n - G) + \gamma \frac{900}{T+273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0,34u_2)} \quad (2.1)$$

donde

ET_0	Evapotranspiración de referencia (mm dia^{-1})
R_n	Radiación neta en la superficie del cultivo ($\text{MJ m}^{-2} \text{ dia}^{-1}$)
R_a	Radiación extraterrestre (mm dia^{-1})
G	Flujo del calor de suelo ($\text{m}^{-2} \text{ dia}^{-1}$)
T	Temperatura media del aire a 2m de altura ($^{\circ}\text{C}$)
U_2	Velocidad del viento a 2 m de altura (m s^{-1})
e_s	Presión de vapor de saturación (kPa)
e_a	Presión real de vapor (kPa)
$e_s - e_a$	Déficit de presión de vapor (kPa)
Δ	Pendiente de la curva de presión de vapor ($\text{kPa } ^{\circ}\text{C}^{-1}$)
γ	Constante psicrométrica ($\text{kPa } ^{\circ}\text{C}^{-1}$)

Esta ecuación utiliza datos del ambiente como son radiación solar, temperatura del aire, velocidad del viento. Estos datos para ser operativos se toman a 2 metros de altura respecto a la superficie.

DATOS

Se puede hacer referencia a datos anuales, mensuales, semanales o diarios para utilizar la ecuación, donde cada frecuencia implica una corrección diaria o semanal de los datos necesarios. Estos datos corresponden a:

- **Localización:** La altura sobre el nivel del mar (m) de la zona para la que se determina la ET_0 y su latitud (grados norte o sur) deben ser especificados. Estos datos son necesarios para ajustar algunos parámetros climáticos al valor medio local de la presión atmosférica (función de la elevación del sitio sobre nivel del mar) y para calcular la radiación extraterrestre (R_a) y, en algunos casos, la duración máxima de la insolación (N).
- **Temperatura:** El uso de la temperatura media del aire en lugar de la temperatura máxima y mínima tiene como resultado una presión de saturación más baja (e_s), y por lo tanto un déficit de presión de vapor también más bajo ($e_s - e_a$), resultante en una estimación más baja de la evapotranspiración de referencia.

- Humedad: En caso de que los valores de presión real de vapor no estén disponibles, estos pueden ser derivados de la humedad relativa máxima y mínima.
- Radiación: La radiación neta diaria (promedio) está expresada en Megajoules por metro cuadrado por día ($MJ m^{-2} dia^{-1}$).
- Viento: Es importante verificar la altura a la cual se mide la velocidad, pues velocidades del viento medidas a diversas alturas sobre la superficie del suelo presentan con seguridad valores diferentes.

El Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias de México realizó un proceso de comparación entre los diferentes métodos para determinar la evapotranspiración y resolvió que el que presentaba la mejor estimación es la ecuación Penman-Monteith ([Chavez-Ramirez et al., 2013](#)).

2.2.6. ECUACIÓN DE HARGREAVES PARA EL CÁLCULO DE LA EVAPOTRANSPIRACIÓN

La fórmula de Hargreaves para evaluar la evapotranspiración Potencial necesita solamente datos de temperaturas y de radiación solar ([Hargreaves and Allen, 2003](#)).

Cuando la radiación solar, humedad relativa o velocidad del viento no son obtenidos es posible estimarlos usando ciertas variables alternativas como las temperaturas máximas, mínimas y otras a través de la ecuación de Hargreaves:

$$ET_0 = 0,0023(T_{mean} + 17,8)(T_{max} - T_{min})^{0,5}R_a$$

En donde los parámetros han sido previamente definidos. Las unidades de la evapotranspiración y de la radiación solar incidente están dadas en milímetros por día, además para el uso de la radiación solar extraterrestre nos remitimos a una tabla en función de la latitud y del mes propuesta por Allen ([Richard G. Allen, 1998](#)), la cual originalmente se presenta en MJulio/m²/día pero para usos en Latinoamérica se presenta en mm/día en la tabla 2.1.

2.3. GESTIÓN DEL CONOCIMIENTO

La gestión del conocimiento es un concepto aplicado a las organizaciones en general en su gran mayoría desarrollado en el sector empresarial que pretende organizar y transferir conocimiento entre la misma, de modo que se pueda usar como un recurso en la organización ([Vanegas, 2015](#)). Permite identificar, agrupar, ordenar y compartir de forma sistemática conocimiento afín a las organizaciones para satisfacer necesidades presentes y futuras, buscando identificar y explotar recursos de conocimiento existentes como adquiridos y para desarrollar oportunidades de crecimiento.

Este proceso comparte su desarrollo con el aprendizaje corporativo buscando:

- Identificar, recoger y organizar el conocimiento existente.
- Facilitar la generación de nuevo conocimiento.
- Iniciar procesos de innovación y desarrollo a través de la reutilización de conocimiento que permitan generar un funcionamiento mejorado del negocio.

Lat	Ene	Feb	Mar	Abr	May	Jul	Ago	Sep	Oct	Nov	Dic
22	17.2	16.4	14.8	12.5	10.4	9.8	11.5	13.8	15.7	16.9	17.4
20	17.1	16.3	14.9	12.8	10.9	10.2	11.8	14.0	15.8	16.8	17.2
18	16.9	16.3	15.1	13.1	11.2	10.6	12.2	14.2	15.8	16.7	17.0
16	16.8	16.3	15.2	13.4	11.6	11.0	12.5	14.4	15.8	16.6	16.8
14	16.6	16.2	15.3	13.6	12.0	11.4	12.8	14.5	15.8	16.4	16.6
12	16.4	16.2	15.4	13.9	12.3	11.8	13.1	14.7	15.8	16.2	16.3
10	16.1	16.0	15.4	14.1	12.7	12.2	13.4	14.8	15.7	16.0	16.1
8	15.9	15.9	15.5	14.3	13.0	12.5	13.6	14.9	15.7	15.8	15.8
6	15.6	15.8	15.5	14.5	13.3	12.9	13.9	15.0	15.6	15.6	15.5
4	15.3	15.6	15.5	14.7	13.6	13.2	14.1	15.1	15.5	15.3	15.2
2	15.1	15.5	15.5	14.9	13.9	13.5	14.4	15.1	15.4	15.1	14.9
0	14.8	15.3	15.5	15.0	14.2	13.8	14.6	15.2	15.3	14.8	14.5

Cuadro 2.1: Radiación solar según latitud. Fuente: Propia adaptada de (Hargreaves and Allen, 2003)

¿QUÉ ES CONOCIMIENTO?

Para poder definir el conocimiento es necesario verlo como parte de un proceso el cual permite por una serie de pasos para entender y abstraer ideas sobre el entorno. Este proceso se conoce como pirámide del conocimiento, en donde cada parte de la pirámide representa un grado mas alto de entendimiento basado en la asimilación y en trabajo con el nivel inferior que ocupa.

Esta pirámide tiene como base datos, que son una representación simbólica, atributo o característica de una entidad que no conlleva un valor semántico pero procesado puede servir de apoyo a la toma de decisiones. El siguiente escalón de la pirámide corresponde a la información definiéndose como un conjunto ordenado de datos que forman un mensaje sobre algún fenómeno. Sobre este escalón se forma el conocimiento definido como la capacidad para actuar y producir resultados como el entendimiento la inteligencia y la razón natural; en la punta de la pirámide se encuentra la sabiduría como el grado mas alto de conocimiento (Enciso, 2011). Esta pirámide es presentada en la figura 2.6.

El conocimiento se clasifica en tácito y explícito. El conocimiento tácito es aquel no codificable, cargado de subjetividades, valores e intuiciones que se adquiere por la experiencia y no por estudios formales, el conocimiento explícito es aquel escrito en manuales, libros, etc., y se adquiere a través de estudios formales.

2.3.1. MODELOS DE GESTIÓN DEL CONOCIMIENTO

La gestión del conocimiento es reconocida como de valor estratégico en la organización y se le vincula un carácter de relevancia, pero dada su complejidad e intangibilidad no es fácil su desarrollo. De esta problemática surgen modelos sobre como debe organizarse, gestionarse y adecuarse el conocimiento en una organización, de ahí la creación de modelos de gestión del conocimiento (Perez and Urbaz, 2016).

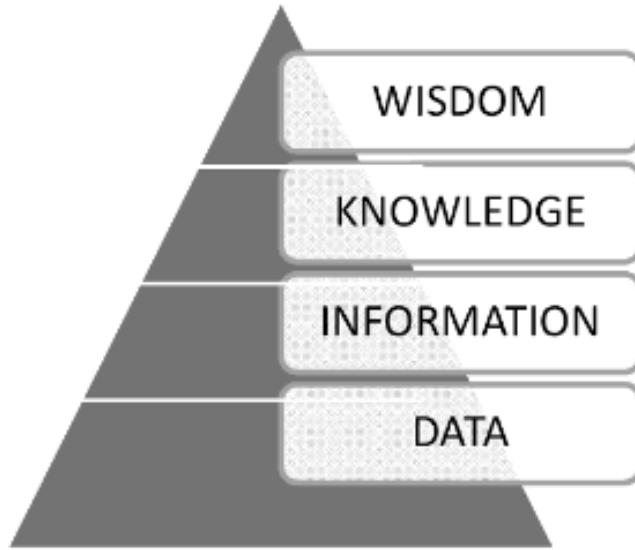


Figura 2.6: Pirámide del conocimiento.
Fuente: Cooplexity (Enciso, 2011)

MODELO DE WIIG

El modelo de Wiig propone que para que el conocimiento pueda ser útil y valioso debe ser organizado. Este se basa en cinco principios básicos que son creación, captura, renovación, compartir y uso del conocimiento. Destaca el uso de las TIC en el proceso de distribución del conocimiento, mas no en los demás procesos como la captura (H.D., 2002).

MODELO DE NOKANA Y TAKEUCHI

En este modelo la gestión del conocimiento abarca los procesos de creación o captación, estructuración, transformación y transferencia del conocimiento, almacenamiento e incorporación en los procesos de la organización. Abarca el conocimiento individual y colectivo. Enfatiza que crear conocimiento no es crear información sino procesarla, hacerla útil para la organización fusionándola con ideas. No destacan el uso de las TIC en la gestión del conocimiento, tan solo se enfocan en los procesos humanos (Diaz, 2005).

MODELO DE INTEGRACIÓN DE TECNOLOGÍA DE KERSHBERG

Kershberg reconoce diferentes medios por los cuales recabar conocimiento y se integran varios componentes que llama *capas*, a saber capa de presentación, capa de gestión del conocimiento y capa de datos. El modelo establece diferentes tecnologías en un alto nivel de integración permitiendo trabajar con estándares comunes (Diaz, 2005).

Este modelo propone:

- Asegurar la calidad de los datos e información.
- Transformación de los datos y la información en conocimiento.

- Almacenamiento eficiente y recuperación de metadatos.
- Promoción a tiempo de la diseminación y la distribución del conocimiento.
- Presentación adaptada del conocimiento.

MODELO INTEGRADO SITUACIONAL DE RIESCO

Es un método definido desde una doble percepción, la holística y la particular, además de una perspectiva social y tecnológica. Llega a las definiciones de conocimiento y su gestión desde una visión situacional. Propone una generación de conocimiento mediada por la memoria corporativa, TIC y redes todas relacionadas entre si consolidando una sola estructura (Quintero et al., 2014). Este modelo integrado situacional se describe en la figura 2.7.

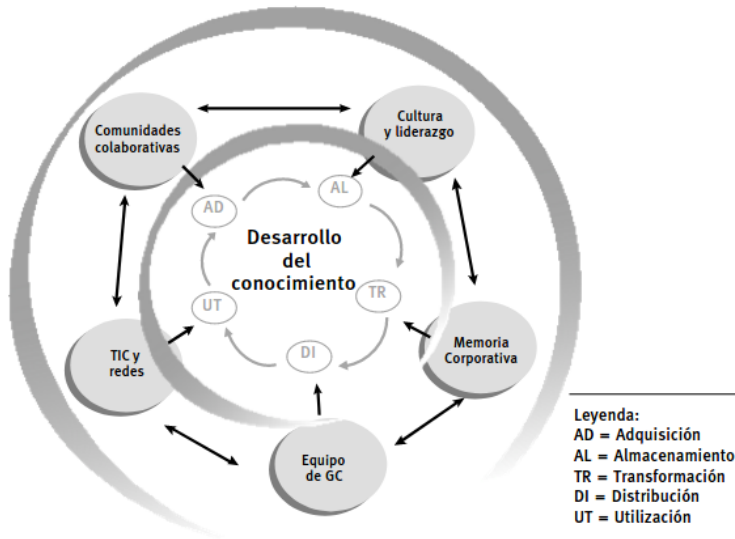


Figura 2.7: Arquitectura del modelo integrado-situacional
Fuente: Adaptado de (Quintero et al., 2014)

MODELO DE GESTIÓN TECNOLÓGICA DEL CONOCIMIENTO DE PANIAGUA Y LÓPEZ

Este se compone de recursos de conocimiento, actividades de transformación del mismo y factores que influyen su gestión. Los recursos los interpreta como el entendimiento de los agentes involucrados, la cultura de la organización, la infraestructura, los artefactos, la estrategia y los recursos externos.

Como actividades de transformación del conocimiento propone la socialización, exteriorización y combinación del mismo. Los factores de influencia los define como aquellos que de alguna forma pueden modificar el comportamiento del conjunto, como la influencia de los agentes, de la gestión y del entorno.

MODELO MONITOR DE ACTIVOS INTANGIBLES DE SVEIBY

Es un modelo basado en la fortaleza de los intangibles que se estructura sobre la diferencia que el creador encuentra entre el valor de la organización en los libros contables y el valor de las acciones en el mercado, indicando que las personas son el único agente verdadero en las organizaciones y las que crean sus estructuras. El monitor de activos intangibles es una herramienta que permite la medición de estos activos.

Los indicadores de este modelo son de crecimiento-innovación que resalta el conocimiento adquirido tanto de forma empírica como de forma académica, los de ciencia refuerzan los de crecimiento al incluir los valores añadidos de las áreas de ventas, la actitud de la gente y personal de apoyo, los cuales generalmente no se consideran activo de la organización además de los indicadores de estabilidad que dan cuenta de cuanto tiempo o cuanta resiliencia puede tener la empresa según el grado de permanencia de los individuos en la organización. Esta garantiza la permanencia de los conocimientos en la empresa (Vanegas, 2015).

2.3.2. DESARROLLO DEL PROCESO DE GESTIÓN DEL CONOCIMIENTO

Con los objetivos planteados por la gestión del conocimiento, el proceso por el cual se desarrolla la misma es a través de la espiral presentada en la figura 2.8 adaptada de (Wesney, 2008).

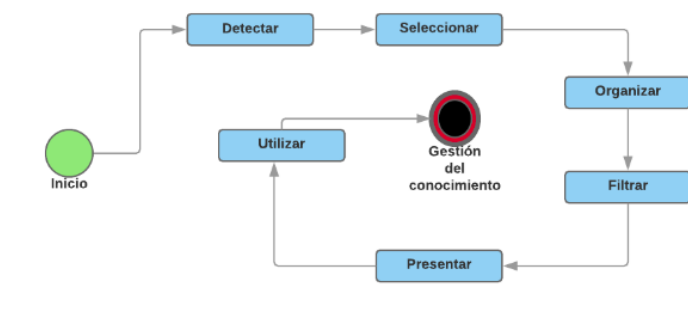


Figura 2.8: Espiral de la gestión del conocimiento
Fuente: Propia a partir de (Wesney, 2008)

De esta manera la forma de implementar un proceso de gestión del conocimiento podría verse así (Torres, 2017):

Detectar: En este proceso se ubican los modelos cognitivos y activos inherentes a la organización buscando determinar nuevas fuentes de conocimiento. Se pueden lograr buscándolos tanto de forma interna como externa.

Seleccionar: Al seleccionar se escoge el modelo de conocimiento frente al cual se realizará la gestión del mismo. Los criterios para seleccionarlo son propios del ejecutor.

Organizar: Es el momento en donde se estructura la representación del modelo. La primera aproximación se da generando ideas del mismo y encontrando patrones para luego codificarlos buscando que puedan ser accedidos por cualquier miembro

de la organización para generar una transferencia del conocimiento ayudado por interfaces de acceso.

Filtrar: Es la etapa del proceso en donde generamos consultas automatizadas para acceder a la información. Estas basadas en estructuras de acceso simple y compleja.

Presentar: Los resultados obtenidos deben ser presentados ya sea a personas o máquinas y debe estar enfocada en abarcar y ampliar el rango de la comprensión humana.

Utilizar: Se precisa en el acto de aplicarlo al problema objeto de resolución buscando generar una retroalimentación cuyo alcance es la generación de valor y la mejora constante.

De esta forma se aprecia que es necesario motivar la conversión del conocimiento de carácter tácito a explícito, elevando su comprensión y uso para que pueda ser abordado por todos los integrantes de la organización, desde el individuo hasta nivel interinstitucional, paso mediado por la transformación del conocimiento, como se ve en la figura 2.9.

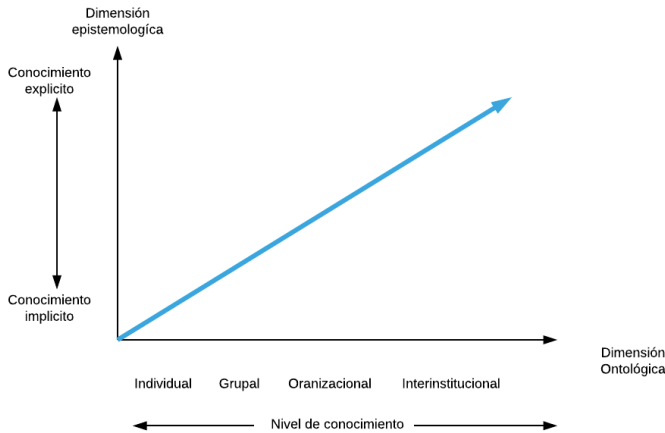


Figura 2.9: Creación y transformación de conocimiento
Fuente: Inteligencia integral de negocios (Torres, 2017)

2.3.3. ETAPAS QUE COMPONEN EL PROCESO DE GESTIÓN DEL CONOCIMIENTO

El proceso de gestión de conocimiento inicia con la identificación de los conocimientos internos y externos disponibles. Los internos son los que provienen de la organización como el conocimiento latente, el periférico, el básico vital, etc.; tanto que los externos se extraen del entorno mediante acuerdos con los creadores de información, fusiones, técnicas de benchmarking, etc.

Luego de esto se selecciona entre las fuentes consultadas los conocimientos asequibles y útiles para la organización considerados activos. Los conocimientos explícitos serán inventariados y puestos a disposición de quien los precise, los conocimientos tácitos serán plasmados en una mapa de localización para identificar las fuentes de los mismos.

Luego se procede a la etapa de creación de conocimiento se utilizan las fuentes anteriores para cruzarlas entre si y con el entorno organizacional para crear nuevo conocimiento y enviarlo a donde es necesario en la organización.

La última fase de este proceso es asimilar y utilizar este conocimiento almacenado y generado (Castillo and Azkuenaga, 2007). El proceso se presenta en la figura 4.19.

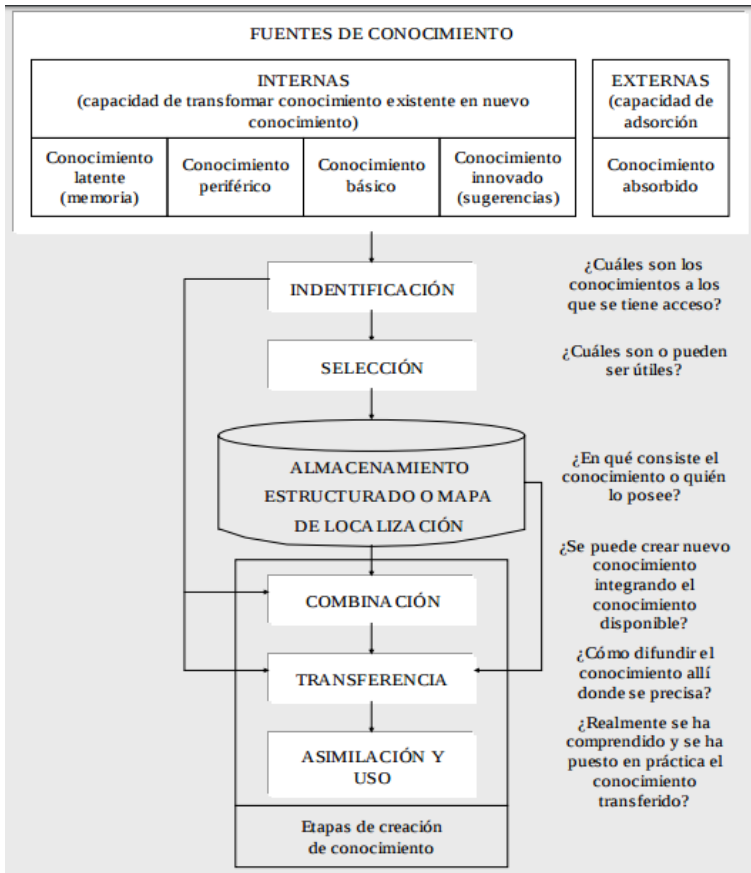


Figura 2.10: Proceso de gestión del conocimiento

Fuente: (Castillo and Azkuenaga, 2007)

2.4. INTELIGENCIA COMPUTACIONAL

La Inteligencia Artificial es una representación de inteligencia provista por las máquinas, donde esta toma el carácter de *inteligente* cuando percibe su entorno de una

forma racional y puede tomar decisiones con respecto a este. Coloquialmente este término es usado cuando un sistema informático o una máquina imita funciones cognitivas de decisión propias de un ser humano como lo son el aprender y resolver problemas (Alpaydm, 2010). Entre las formas que estos se pueden clasificar están:

Inteligencia Artificial convencional También llamada simbólico-deductiva que fundamenta sus análisis en el desarrollo formal y estadístico. Presenta herramientas como el razonamiento basado en casos, los sistemas expertos, las redes bayesianas, la inteligencia artificial basada en comportamientos y la administración de procesos inteligentes.

Inteligencia Computacional Centra su estudio en mecanismos adaptativos buscando no analizarlos por medio de procesos heurísticos proponiendo alternativas como redes neuronales, computación evolutiva, inteligencia de enjambre, sistemas inmunes artificiales o lógica difusa (Poole et al., 2006).

Con base a esto la inteligencia computacional es programar computadores para mejorar un criterio de rendimiento usando datos de ejemplo o experiencia anterior. El modelo puede ser predictivo para hacer estimaciones de comportamientos futuros o descriptivo para ganar conocimiento de los datos, o ambos.

2.4.1. TIPOS DE APRENDIZAJE

Los métodos que abarca la inteligencia computacional para resolver problemas mediante la clasificación o predicción son muy variados y su uso depende en gran medida de los datos disponibles para *enseñarle* el problema al sistema. La forma en que el sistema asimila da origen a los diversos tipos de aprender que puede tener, estos pueden ser supervisados o no supervisados.

En el aprendizaje no supervisado tenemos patrones de entrada al sistema pero no ayudas de aprendizaje, así que la solución tiene parámetros de entrada e intenta identificar patrones entre ellos para clasificarlas en categorías similares. Esto se hace con un conjunto de datos de entrenamiento, con los cuales le decimos a nuestro sistema la forma en que le ingresarán datos.

El aprendizaje supervisado provee patrones de entrenamiento con una salida deseada, siendo así el sistema puede configurarse para con las entradas provistas alcanzar las salidas esperadas en cada paso de su entrenamiento, e ir variando su arquitectura interna para cada vez acercarse más a la salida deseada (Kriesel, 2007).

2.4.2. HERRAMIENTAS DE INTELIGENCIA COMPUTACIONAL REDES NEURONALES

Una neurona es la estructura mínima de almacenamiento del cerebro humano, en un intento por emular el comportamiento de este de forma computacional se creó una neurona digital que copia la estructura básica de una real teniendo diferentes señales de entrada donde cada una se asocia a un peso que determina la relevancia de esa entrada en el sistema. Todas estas entradas se suman y son llevadas a una función de activación la cual pondera la señal y realiza la operación correspondiente cuando esta señal supera un umbral de trabajo. Esta disposición se presenta en la figura 2.11.

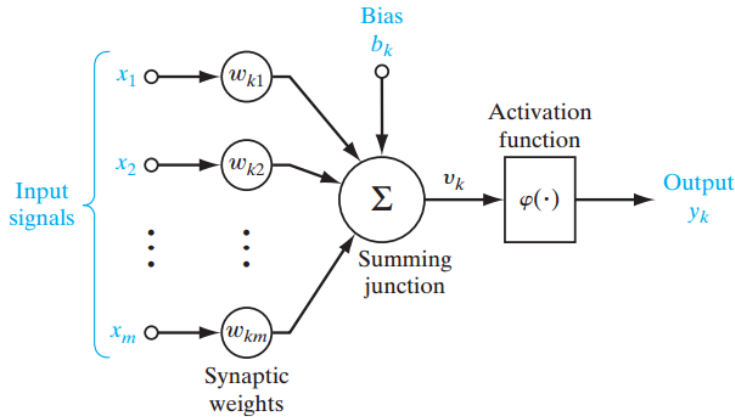


Figura 2.11: Representación digital de una neurona
Fuente: (Haykin, 2009)

Una red neuronal es un procesador masivo paralelo distribuido hecho de una unidad simple de procesamiento llamado neurona que busca imitar el funcionamiento biológico de las neuronas humanas, tiene una disposición natural a almacenar experiencias de conocimiento y presentarlas para su uso (Haykin, 2009). Se hace similar al cerebro humano en dos formas:

- El conocimiento es adquirido por la red desde el medio ambiente a través de un proceso de aprendizaje.
- Una conexión entre neuronas fuerte, conocido como pesos sinápticos son usados para almacenar el conocimiento adquirido.

Las redes neuronales cuentan con las siguientes propiedades y capacidades (Haykin, 2009):

No linealidad Una neurona artificial puede ser lineal o no lineal, una red neuronal que incluya una neurona no lineal es de hecho no lineal

Trazado de entrada a salida El aprendizaje supervisado involucra la modificación de los pesos sinápticos de una neurona al aplicar un conjunto de entrenamiento y en cada etapa de este cada entrada esta asociada a una salida, así que la red puede hacer un trazado completo entre las entradas propuestas y las salidas esperadas.

Adaptabilidad Las redes neuronales tienen la capacidad de cambiar sus pesos sinápticos para adecuarse al ambiente que las rodea. En particular una red neuronal entrenada para trabajar en un ambiente puede fácilmente re-entrenarse para adaptarse a otro pudiendo cambiar sus pesos sinápticos en tiempo real.

Evidencia de la respuesta En el contexto de clasificación de patrones una red puede ser diseñada para seleccionar un patrón y también para presentar el grado de confianza de su selección. Esto se hace para mejorar el rendimiento de clasificación de la red.

Información en contexto El conocimiento es representado en la estructura y activación de la red. Cada neurona afecta potencialmente a las demás.

Tolerancia a fallos Una red neuronal implementada en hardware es capaz de una computación robusta ya que puede perder parte de su configuración si sufre algún ataque o daño pero el conocimiento adquirido se encuentra en toda la red lo que la hace descentralizada.

MAQUINAS DE SOPORTE VECTORIALES(SMV)

Son un conjunto de algoritmos de aprendizaje supervisado que se enfocan en clasificación y regresión, se pueden representar como la unión de puntos en un espacio muestral separando los dos mas representativos mediante un hiper-plano llamado vector soporte, también presentado en la literatura como *kernel machine* que debe su popularidad de uso a (Alpaydm, 2010):

- Es un discriminante basado en el principio de Vapnik's que busca nunca solucionar un problema complejo aumentando su complejidad, buscando en la solución calcular los límites de esta antes que los parámetros asociados con la misma.
- Después del entrenamiento el parámetro del modelo lineal vector de peso puede ser escrito en término de un subconjunto del conjunto de entrenamiento llamado *vector soporte*. Estos números también son usados para estimar el error general.
- La salida es escrita como la suma de los vectores con mas peso y son llamadas *funciones de kernel* que son medidas relacionadas con los parámetros de similitud entre las instancias de datos.
- En la mayoría de algoritmos de aprendizaje los datos son representados como vectores y se pueden realizar operaciones como el producto punto (en el caso de las redes neuronales) o distancia Euclidiana. Una función kernel permite ir mas allá, por ejemplo G_1 y G_2 pueden ser gráficas y $K(G_1, G_2)$ la cantidad de caminos compartidos que se puede calcular sin necesidad expresar G_1 y G_2 como vectores.
- Los algoritmos basados en kernel son formulados como problemas de optimización convexa y esta es una forma de optimizarlos analíticamente pues no es necesario abordarlos con métodos heurísticos, inicializaciones, verificando convergencia y herramientas así.

Estas características se aprecian al momento de realizar un proceso de clasificación con una SMV, en donde un hiper-plano identifica características en los datos y los separa según estas. Un ejemplo de aprecia en la figura 2.12.

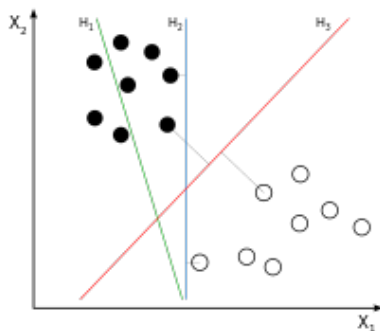


Figura 2.12: Proceso de clasificación con SMV

Fuente:(Alpaydın, 2010)

LÓGICA DIFUSA

La lógica difusa brinda una herramienta de inferencia que permite simular los procesos del razonamiento humano basado en conocimiento y su teoría parte de una base matemática que incorpora la incertidumbre de los procesos cognitivos para poderlos manejar de forma computacional (Morcillo, 2014).

La incertidumbre en las abstracciones que se realizan del mundo viene cuando no es posible distinguir un estado del siguiente, por ejemplo la expresión *El agua debe estar ligeramente caliente* para un ser humano tiene una connotación y un contexto pero no es posible que un sistema informático tenga asociado un valor a la expresión *ligeramente* pudiendo ser según interpretación una temperatura u otra, a esto lo llamamos incertidumbre.

Básicamente la lógica difusa es una lógica multivariada que permite representar matemáticamente la incertidumbre y la vaguedad proporcionando herramientas formales para su tratamiento. Se diferencia de la probabilidad en que esta representa información sobre la frecuencia de ocurrencia de un suceso basado en el total de eventos posibles, en cambio la lógica difusa indica el grado de pertenencia de un evento con respecto a otro donde las propiedades de ambos no están definidas de manera precisa (Duarte, 1999).

La lógica difusa plantea que la representación de un problema complejo no puede ser dada de forma precisa y propone el empleo de la teoría de conjuntos difusos para su representación y así describir las relaciones entrada salida mediante proposiciones condicionales del tipo *SI-ENTONCES*, por ejemplo, *SI la lluvia es muy fuerte ENTONCES disminuye el caudal de riego*). Luego el uso de un método de inferencia generalizado conocido como reglas de composición de inferencia para combinar la información y generar acciones dando pie a las siguientes características:

- El razonamiento **exacto** se considera un caso particular del aproximado, así pues cualquier sistema formal puede ser *fuzzificado* y se puede representar el conocimiento humano de forma sistémica para su inclusión en modelos de ingeniería.
- El conocimiento se interpreta como una serie de reglas o **restricciones difusas** sobre las variables y su fortaleza se evidencia cuando un modelo exacto es difícil de obtener.

- La inferencia puede verse como un proceso de propagación de estas restricciones difusas.
- Se utiliza ampliamente en sistemas de ayuda a la decisión. La lógica difusa permite obtener decisiones con valores incompletos o información incierta.

Los sistemas difusos son recomendables en aquellos problemas complejos donde no existe un modelo matemático simple asociado. Igualmente en procesos que obedecen a un comportamiento no lineal, la solución difusa plantea ventajas. Requiere que el conocimiento experto sea expresado lingüísticamente; la representación gráfica de los conjuntos se presenta en la figura 2.13.

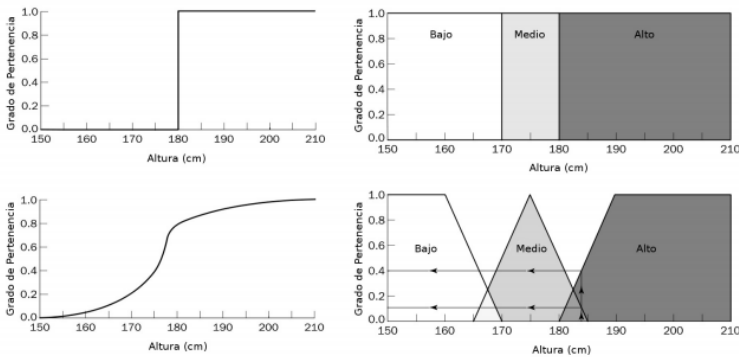


Figura 2.13: Representación gráfica de un conjunto difuso
Fuente:(Morcillo, 2014)

2.5. EL PARADIGMA IOT

El IoT por sus siglas en inglés es un paradigma que abarca diversos tipos de tecnologías y diversas configuraciones de las mismas para obtener datos sobre el entorno en donde es instalada y convertir estos en información útil a fines específicos, el cual consta generalmente de tres capas. Una capa de sensores permite levantar datos de interés, una capa de comunicaciones brinda transporte de esos datos a un sitio donde puedan ser almacenados en donde otra capa analiza y extrae información de ellos.

La visión principal del IoT es la creación de un mundo inteligente donde lo real, digital y lo virtual converjan para crear un entorno que proporcione inteligencia a la energía, salud, transporte, ciudades, industria, edificios y demás áreas de la vida (Bagula, 2015).

Diversas organizaciones a nivel mundial han definido el concepto de Internet de las Cosas desde su visión nutriendo el desarrollo teórico de este ecosistema, abarcan principios como comunicaciones máquina a máquina, redes de sensores inalámbricos entre otros. Algunas entidades que han definido IoT son (Zhou, 2013):

CASAGRAS (Coordinación y soporte de actividades y estandarización con RFID): IoT es una infraestructura de red global conectada de forma física y virtual con objetos a través de la captura de datos y capacidades de comunicación. Esta infraestructura

abarca Internet y despliegues de red independientes ofreciendo identificación de objetos particulares además capacidad de sensado y comunicación de los mismos caracterizado por un grado alto de autonomía al momento de captura y transferencia de datos.

2

SAP(Análisis de Sistemas y Desarrollo de Programación): IoT está creando un mundo donde los objetos físicos se integran en redes de información y donde pueden ser participantes activos en procesos de negocio. Los servicios están disponibles para actuar con estos objetos cambiando su estado e información a través de la apropiada seguridad.

EPoSS's (Plataforma de Tecnología Europea en Integración de Sistemas Inteligentes): IoT es la red formada por objetos o cosas que tienen identidades, personalidades virtuales operando en espacios inteligentes usando interfaces para conectarse y comunicarse con usuarios, redes sociales y contextos ambientales.

CERP's(Cluster de Proyectos RFID Europeos): IoT es parte íntegra del futuro de Internet y puede ser definida como una infraestructura de red global dinámica auto-configurable basada en estándares y protocolos de comunicación donde los objetos físicos y virtuales tienen atributos, personalidades e interfaces inteligentes añadiendo información a la red de forma continua.

APLICACIONES DE IOT

Varios centros de investigación a lo largo del mundo han presentado su perspectiva de las diferentes áreas de aplicación del IoT y el impacto que tienen, algunas siendo ejercicios de prospectiva sin mas base que la intuición de los investigadores, otros estudios se basan en proyecciones de las consideraciones actuales del mercado, siendo un desarrollo de perspectiva en el corto plazo.

El laboratorio de tecnología e investigación en redes emergentes y sistemas distribuidos (TRENDS por sus siglas en inglés) adscrito a la Universidad de Tecnologías de la Información en Pakistan presenta un análisis de áreas de interés del IoT con base en los proyectos de investigación y desarrollo actuales, presentado en la figura 2.14.

En esta se aprecia que el interés de los investigadores se centra en la modernización de la industria y el desarrollo de ciudades inteligentes además de trabajo sobre energía e interconexión de vehículos. El área de interés del proyecto en agricultura inteligente es el sexto en interés con un desarrollo mayor en la América que en otra zona geográfica. Este estudio relacionó 640 iniciativas a nivel global.

La Universidad en Tecnologías de la Información en Pakistan presenta las empresas que están cambiando el modo de desarrollo y operación de las granjas llevándolas al concepto de granja inteligente. Esta información se presenta en la figura 2.15 en donde se muestra que existen empresas trabajando en riego inteligente evidenciando el interés por este tipo de soluciones; otras áreas de trabajo a nivel comercial son el software para administrar granjas, la agricultura de precisión, el desarrollo de mercados, la robótica y drones, manejo y desarrollo de sensores, sistemas de analítica de plantas, levantamiento y manejo de datos de animales y granjas de próxima generación.

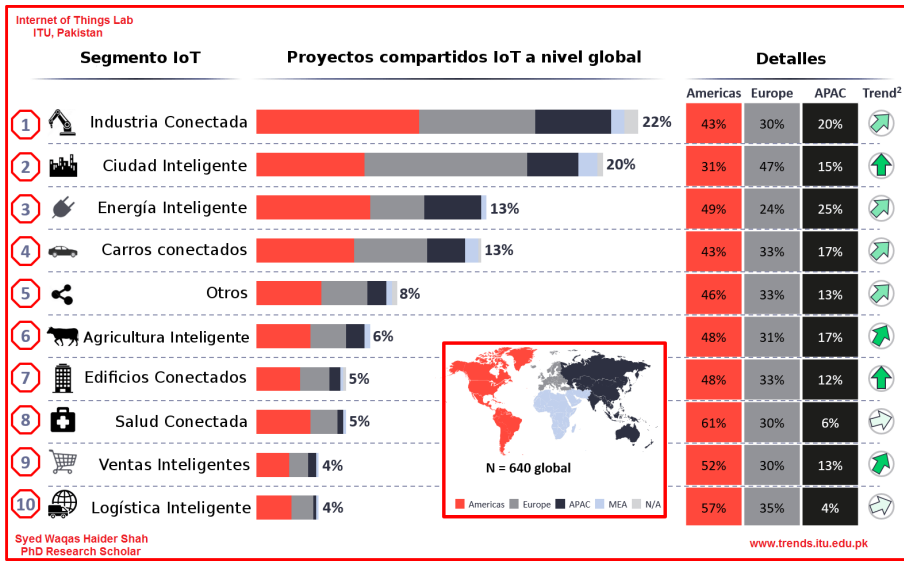


Figura 2.14: Proyectos en desarrollo en IoT
Fuente:(Khan et al., 2012)

2.5.1. PROTOCOLOS DE COMUNICACIÓN IOT

Los sistemas embebidos utilizan diversos protocolos de comunicación para transportar información de sensores y luego estos se unen con otros en Internet. Los Protocolos de Internet (IPs) no son nuevos, pero los IPs dedicados a IoT si lo son y hay múltiples capas de IP sobre sockets TCP-IP cada una con ventajas y desventajas, siendo ellos tan diferentes que se hace necesario conocer sus diferencias para escoger el tipo de protocolo IoT a usar en cualquier aplicación desarrollada según requerimientos en ancho de banda, rendimiento en tiempo real y memoria. Estos son:

- De Consumo vs Industrial
- Servicios Web
- Servicios IoT
- Publicación / Suscripción
- Petición / Respuesta

INTERNET

El Internet es la suma de la infraestructura usada para darle una ruta a un paquete IP desde una fuente a un destino. La World Wide Web por comparación es un sistema de aplicaciones que corre sobre Internet, es una herramienta construida por personas para el intercambio de información y en los últimos 20 años ha sido desarrollada y redefinida

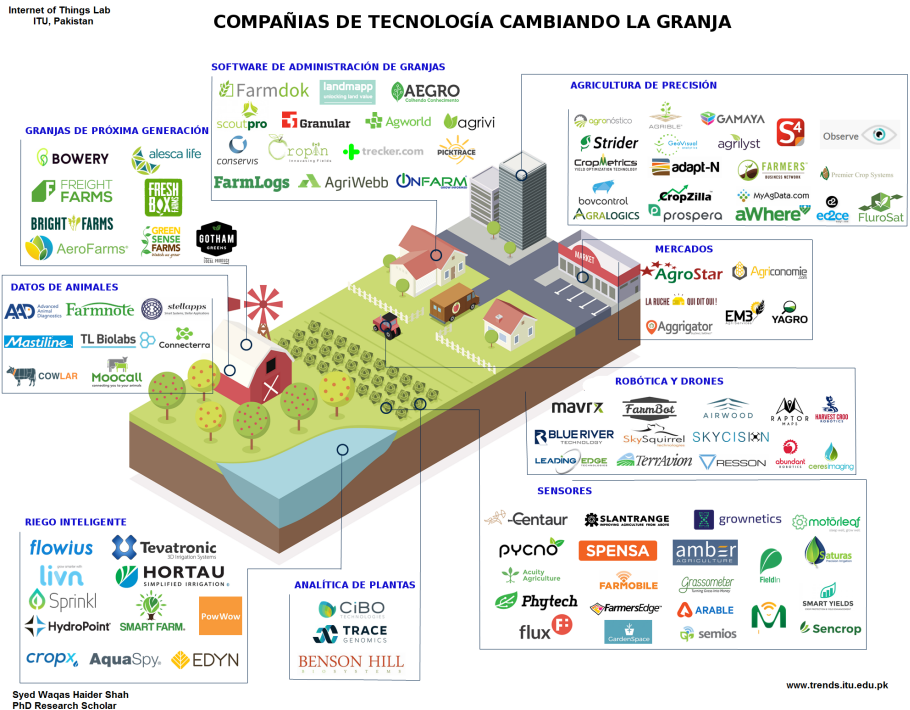


Figura 2.15: Empresas en agricultura inteligente
Fuente: TRENDS, (Khan et al., 2012)

de forma tal que es usada de forma fácil y productiva por usuarios no técnicos. Por ejemplo, la interfaz de Internet ahora incluye email, buscadores, navegadores, redes sociales como Twitter, Facebook, etc.

Por comparación en el IoT la idea es intercambiar información de la misma manera sobre Internet, pero no se cuenta con el equivalente de navegadores o medios sociales para compartir la información. No hay consensos sobre la forma de hacerlo debido a la diferencia de IoT con Internet en cuanto a velocidad y escalabilidad aunque se han hecho intentos. Muchas de las visiones sobre el uso de estos servicios en IoT podrían o deberían chocar entre ellos.

PILA TCP/IP

La pila del protocolo TCP/IP es el corazón del Internet y la Web, puede ser representada por siete capas en el modelo de abierto de interconexión de sistemas(OSI). Dentro del paradigma IoT resultan de interés tres de estas siete capas, la capa física, la de red y la de transporte (Zhou, 2013).

Capa física Los mas comunes protocolos de capa física usados en los sistemas embebidos son:

- Ethernet (10 Mbps, 100Mbps, 1Gbps)
- Wifi (802.11 b/g/n)
- Protocolo Serial Punto-a-Punto (PPP)
- GSM, 3G, 4G, LTE

Capa de red Es donde el Internet vive, es nombrado así porque interconecta redes, entre las capas físicas. Aquí es donde ubicamos las direcciones IP.

Capa de transporte Encima de IP tenemos el protocolo de control en transmisión (TCP) y el de Datagrama de Usuario (UDP) que son protocolos de transporte. TCP se usa para nuestras interacciones con la Web y provee una conexión lógica, conocimiento de los paquetes transmitidos, retransmisión de paquetes perdidos y control de flujo; estas características lo hacen un protocolo robusto pero excesivo para un sistema embebido. UDP ha sido usado en servicios de red como DNS y DHCP encontrando ahora un lugar en la transmisión de información de sensores y control remoto. UDP es preferible a TCP en datos en tiempo real como vídeo y sonido.

PROTOSCOLOS IOT

Es posible construir un sistema IoT con los protocolos existentes en la Web incluso si no son eficientes con respecto a los nuevos protocolos, HTTP y Web Sockets son estándares comunes junto a XML o JSON en la carga de información. Cuando usamos un estándar web HTTP JSON provee una capa de abstracción a desarrolladores para crear una aplicación persistente.

HTTP

Es la base del modelo cliente-servidor, para incluirlo en un dispositivo basta incluir un cliente HTTP, no todo un servidor. Es seguro que un dispositivo IoT pueda iniciar conexiones a la web, pero no es seguro que la web pueda iniciar peticiones de conexión cuando los dispositivos IP están conectados lo que permitiría acceso a nuestra red local a través de un dispositivo IoT.

WEBSOCKET

Es un protocolo que provee comunicación duplex sobre una conexión TCP sencilla en donde los mensajes pueden ser enviados entre cliente y servidor, siendo parte de la especificación HTML 5. El estándar WebSocket simplifica la complejidad alrededor de comunicaciones Web y manejo de conexiones (Zanella et al., 2014).

XMPP

Este es un buen ejemplo de una tecnología Web encontrando nuevos usos en el mundo IoT. XMPP tiene sus raíces en la mensajería instantánea y se ha expandido hacia llamadas de audio y vídeo, colaboración, middleware ligeros, sindicación de contenido y enrutamiento de datos XML. La fuerza de XMPP es el direccionamiento, seguridad y escalabilidad; esto lo hace ideal para aplicaciones orientadas a IoT.

CoAP

A pesar de que la Web esta disponible para ser usada en soluciones IoT, esta es muy pesada para los diferentes tipos de servicios que se proyectan ofrecer. En julio de 2013 la Fuerza de Trabajo de Ingenieros en Internet (IETF por sus siglas en inglés) desarrollo CoAP para ser usado en redes y nodos de bajo consumo, CoAP como HTTP es un RESTful (la habilidad de manipular fuentes y recursos a través de un API), semánticamente esta alineado a HTTP. CoAP usa UDP por lo que ciertas funcionalidades de HTTP se replican en CoAP de forma transparente distinguiendo entre mensajes confirmados y no confirmados. Las peticiones y respuestas son manejadas de forma asincrónica. Todos los encabezados, métodos y códigos de estado son codificados lo que reduce la parte superior del protocolo, sin embargo necesita un analizador de protocolo para identificar problemas en la red (Zhou, 2013).

MQTT

Message Queue Telemetry Transport (MQTT) o Telemetría de Transporte de Mensajes en Cola es un protocolo abierto desarrollado y optimizado para dispositivos con bajo ancho de banda, alta latencia o redes poco confiables. Es un protocolo de publicación/suscripción extremadamente ligero ideal para conectar dispositivos a redes con mínimo ancho de banda, usando TCP de base. El objetivo de MQTT son grandes redes de pequeños dispositivos que necesitan ser controlados desde un servidor, no esta diseñado para transferencias dispositivo-a-dispositivo ni para multicast enviando datos a múltiples receptores MQTT es simple y brinda herramientas de control a dispositivos individuales.

El protocolo prevee telemetría y a sido usado por Facebook en su aplicación de mensajería por el bajo ancho de banda requerido además de las siguientes ventajas:

- Es código abierto libre de pagos adicionales y fácil de adaptar.
- Sigue una distribución de información uno a muchos.
- Pequeñas cabeceras de mensaje.
- Múltiples niveles de calidad de servicio.
- Mensajes de comando simples.
- Agnóstico en tipo de datos.
- Conexiones durables.

El protocolo MQTT se basa en publicación/suscripción, es decir, una serie de dispositivos se vinculan/publican a un Broker (corredor como en la bolsa, referido a transacciones) que guarda su información y en la etapa de publicación va a estar disponible el registro histórico de esta hasta que alguien o algún dispositivo diferente se suscriba al Broker para recibirla. En la figura 2.16 se presenta la arquitectura de una comunicación sobre el protocolo MQTT en donde la información se maneja de forma bidireccional por el Broker. El sensor de temperatura publica los resultados de su medición en una tabla

dentro del broker con el t pico "temperatura"; ahora cualquier dispositivo puede suscribirse a este t pico en el broker y consultar la informaci n de la tabla, de esta forma el broker es un intermediario entre los sensores y los dispositivos que quieren tener los datos generados por los mismos.

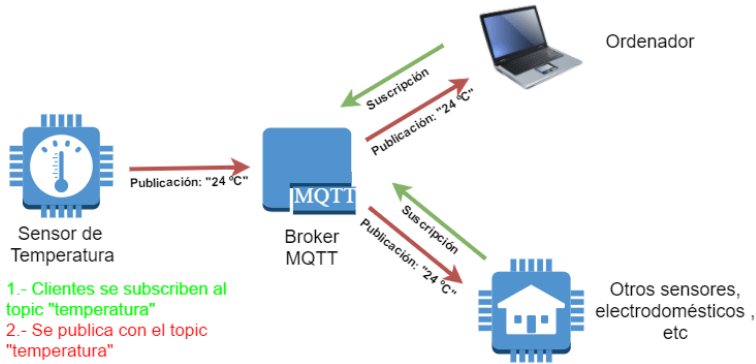


Figura 2.16: Esquema protocolo MQTT
Fuente: Organizaci n MQTT (MQTT.ORG, 2017)

MQTT tiene diferentes comprobaciones para definir la calidad del servicio, si se configura en QoS 0 no se espera una respuesta del receptor as  que un dato se pierde si esta fuera de l nea el receptor o se cae el servidor. En QoS 1 se espera respuesta de verificaci n de recepci n de mensaje sin embargo hay posibilidades de un repetir el dato enviado. En QoS 2 se verifica el dato enviado y su integridad, asegurando que sea el mensaje correcto enviado (MQTT.ORG, 2017).

2.5.2. COMPUTACI N EN LA NIEBLA (FOG COMPUTING)

Fog computing o computaci n en la niebla es una arquitectura de sistema horizontal proyectada para distribuir recursos y servicios de computaci n, almacenamiento, control y redes como un nivel intermedio entre la nube y los objetos ([Group, 2016](#)). Se puede apreciar su descripci n gr fica en la figura 2.17. Es una:

Arquitectura horizontal: Soporta varias empresas verticales y dominios de aplicaci n entregando servicios e inteligencia a usuarios y negocios.

Servicios continuos de la Nube-a-Objetos: Activa servicios y aplicaciones para ser distribuidos cerca de los objetos en IoT y en cualquier parte entre la nube y los objetos.

Nivel de sistema: Extendida desde los objetos sobre los limites de la red a trav s de la nube y sobre m ltiples capas de protocolos(No solamente sistemas de radio).

Se proyecta un desplazamiento del 45% de los datos generados en el 2025 cerca del l mite de la red, la niebla es una arquitectura que resiste grandes cantidades de datos provenientes de IoT, 5G e inteligencia artificial para sistemas cuyas necesidades incluyan seguridad, cognici n(Aut noma centrada en el cliente), agilidad, latencia y eficiencia.

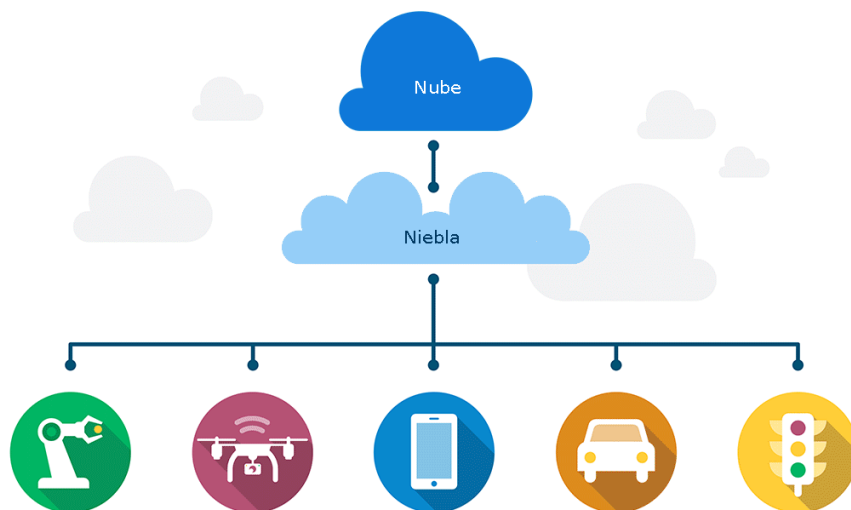


Figura 2.17: Representación de la niebla en Fog Computing
Fuente: OpenFog Consortium (Group, 2016)

La computación en la niebla es propuesta y mantenida por el OpenFog Consortium el cual es fundado por ARM, Cisco, Dell, Intel, Microsoft y la Universidad de Princeton en noviembre de 2015. Esta arquitectura es diseñada para ayudar a líderes de negocio, diseñadores de software, arquitectos de hardware y diseñadores de sistemas para crear herramientas y dispositivos acordes a las necesidades de el creciente mercado IoT.

La arquitectura propuesta para el trabajo bajo concepto de computación en la niebla esta presentada en la figura 2.18 donde se aprecia la interacción de los diversos componentes de la arquitectura. En la capa superior se encuentra la nube o almacenamiento en servidores, interactuando con ella están enrutadores los cuales acopian y direccionan la información hacia la nube en su capa final; a ellos se comunican diversos dispositivos de acceso encargados de suministrar información del vecindario de los objetos y para finalizar los componentes de comunicación están los gateway que nos permiten tomar información en la capa de objetos para su transporte y análisis.

2.5.3. COMPUTACIÓN EN EL TECHO (ROOF COMPUTING)

El proyecto de computación en el techo o Roof Computing por sus siglas en inglés *Standard for an Architectural Framework for Real-Time Onsite Operations Facilitation (ROOF) for the Internet of Things* o Estandar de marco de trabajo para Facilitar Operaciones de Sitio en Tiempo-Real para Internet de las Cosas es una iniciativa reciente iniciada por el doctor Syam Madanapalli en India y apropiada como prototipo de estándar por el Institute of Electrical and Electronics Engineers (IEEE) bajo el código de proyecto 1931.1 actualmente en estado activo el cual busca definir funciones de interoperabilidad técnica de sistemas IoT a nivel local (Madanapalli, 2017). La representación gráfica de donde se ubica la computación de techo se presenta en la figura 2.19.

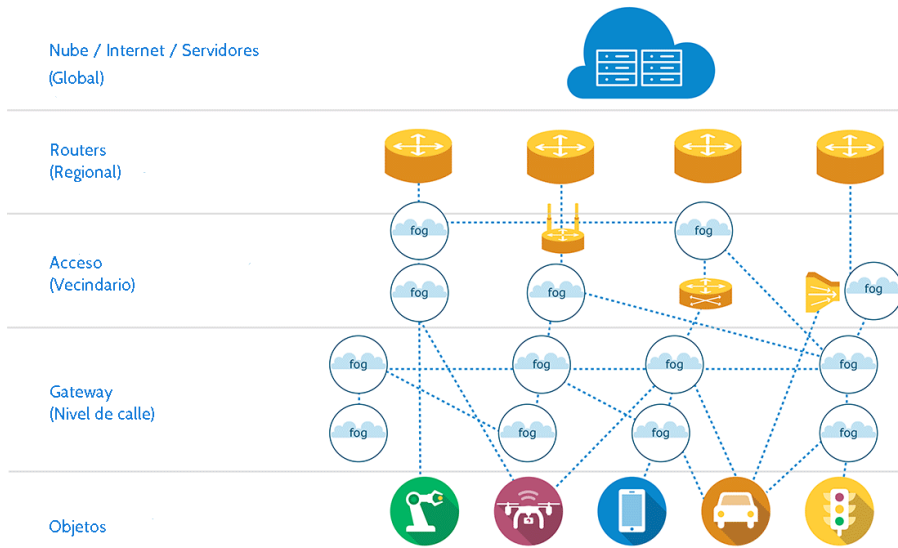


Figura 2.18: Arquitectura computación en la niebla
Fuente: OpenFog Consortium (Group, 2016)

En este contexto las aplicaciones de IoT deben operar con cierta autonomía en ciertos periodos de tiempo con el contexto local sea una casa, fabrica, oficina etc. Para conectarse a la niebla y luego a la nube solo cuando sea necesario o en búsqueda de algún valor añadido.

Al ser una capa adicional de trabajo con una funcionalidad específica reduce la complejidad de la implementación, incrementa la seguridad y brinda agilidad en el manejo de redes IoT de gran envergadura. Esto basado en cuatro funcionalidades:

- Conectividad e interoperabilidad.
- Contexto de trabajo y toma de decisiones en tiempo real.
- Manejo de información y conectividad eficiente a la nube.
- Seguridad y privacidad por diseño.

Estas funcionalidades se presentan en la figura 2.20. Según el paradigma el IoT necesita dos formas de analítica, una de contexto y otra de grandes datos (Big Data), la primera basada en la interpretación en tiempo real de las acciones y las respuestas a esas mismas, la segunda para generar modelos predictivos en el largo plazo.

El IoT genera datos permanentemente la cual no siempre tiene sentido o relevancia para la acción concreta, para esto el techo filtra los datos no deseados y transmite solamente la información útil para futuros análisis incrementando la eficiencia del canal de comunicaciones, pudiendo almacenar información de forma local si este canal se pierde para ser reenviada cuando se logre la conexión.



Figura 2.19: Ubicación de Roof Computing
Fuente:IEEE Standard Association (Madanapalli, 2017)

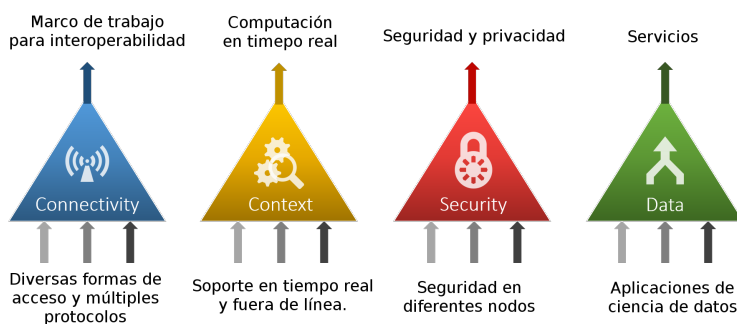


Figura 2.20: Cuatro funcionalidades Roof Computing
Fuente:IEEE Standard Association (Madanapalli, 2017)

2.5.4. CRITERIOS PARA REALIZAR LA IMPLEMENTACIÓN DE UN SISTEMA IOT

Considerando la particularidad del IoT su implementación práctica no se puede realizar con un sistema genérico, es necesario desplegar una estrategia para realizar un proyecto de despliegue exitoso (KaaIoT, 2017). Aún así es posible conseguir un proyectos de forma rápida y sostenibles a largo plazo asegurando la calidad del mismo teniendo en cuenta los pasos de implementación consignados en la figura 2.21.

Donde podemos definir los pasos de la metodología al momento de realizar la implementación.

Definir los casos de uso prioritarios: Es necesario empezar con una clara estrategia de negocios sobre los casos de uso que se implementaran y cuales son de importancia en el largo plazo. Una buena proyección al inicio del desarrollo salvara costos de implementación innecesarios, al mismo tiempo una clara estrategia a largo plazo minimiza riesgos de limitaciones posteriores.

Perfilar los activos tecnológicos clave: Un conocimiento a profundidad de la finalidad

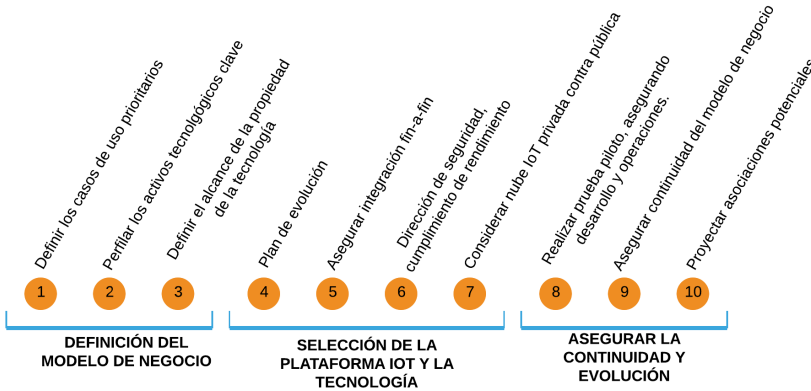


Figura 2.21: Secuencia de implementación de un proyecto IoT
Fuente: Adaptado de (KaaloT, 2017)

en implementación es necesaria para sopesar el uso de sensores, plataformas IoT, hardware, dispositivos de usuarios finales, infraestructura de hospedaje, procesamiento de datos y analítica. Es necesario asegurar que los elementos en la lista son interoperables e integrables.

Definir el alcance de la propiedad de la tecnología: Generalmente es estratégico para un negocio IoT mantener la propiedad sobre la tecnología IoT, pero dependiendo del proyecto se puede desear la libertad y flexibilidad de una plataforma de código abierto o una comercial, tanto en software como en hardware. Es posible cambiar estas decisiones en las fases siguientes del proyecto pero vendrán con costos de transferencia de tecnología, ciclos de desarrollo adicionales y potenciales quiebres en el flujo de información mientras se migra el sistema.

Plan de evolución: Cuando el modelo de negocio ha tomado forma es tiempo de seleccionar una plataforma IoT asegurando que los casos de uso prioritarios son completamente soportados y las funcionalidades puedan ser implementadas en un razonable tiempo. Ya que una plataforma IoT es el pilar fundamental en el proyecto es el primero en la lista, pero el hardware o la solución de analítica necesitan ser escogidas con la misma rigurosidad.

Asegurar integración fin-a-fin: Para asegurar la interoperabilidad se debe asegurar que la plataforma IoT es compatible con los diferentes tipos de hardware, sensores, gateways y soportar los tipos de protocolos de comunicación que se van a necesitar, asegurando que la API de desarrollo este disponible para un uso abierto y que estas funcionalidades corresponden a las que va a necesitar el usuario final.

Dirección de seguridad, cumplimiento de rendimiento: Dependiendo del ambiente de negocio un alto nivel de seguridad y rendimiento alto son dos prerrequisitos para

una implementación exitosa, seleccionando un sistema que brinde la capacidad de encriptar la información en caso de ser necesario y el rendimiento avalado por una alta disponibilidad de servicio, escalabilidad elástica y soporte a nuevos desarrollos.

Considerar nube IoT, privada contra pública: La tecnología IoT se encuentra disponible por varios distribuidores los cuales ofrecen un acceso en la nube con plataforma privativa, un acceso en la nube con una plataforma abierta o ambas opciones implementadas de forma local, teniendo en cuenta que la plataforma IoT es un activo estratégico para el proyecto debe ser motivo de análisis este ítem. Bastantes proyectos usan un modelo híbrido para ubicar el mejor desarrollo.

Realizar prueba piloto, asegurando desarrollo y operaciones: Ir realizando pruebas piloto de los diferentes módulos de la tecnología IoT es una práctica recomendada para realizar pruebas tempranas de velocidad, desarrollo y operaciones, usabilidad entre otras.

Asegurar continuidad del modelo de negocio: Para lograr un desarrollo exitoso en el largo plazo la estrategia IoT debe ser reajustada constantemente, esto con el fin de realizar una mejora constante del tipo de arquitectura para expandirla, todos los días se desarrollan nuevos contenedores, arquitecturas de micro-servicios y demás componentes que se deben tener en cuenta para una correcta actualización de servicios.

Proyectar asociaciones potenciales: Los socios pueden dar mayor velocidad a las iniciativas IoT y a la cadena de valor creada por medio del proceso tecnológico.

2.5.5. PLATAFORMAS IOT

Existen plataformas que permiten integrar las diferentes capas pertenecientes al paradigma IoT las cuales cuentan con diferentes alternativas de uso, algunas se presentan a continuación.

Lumada IoT Platform Es una plataforma perteneciente a la empresa Hitachivantara, el código es cerrado y cuentan con un core, una sección de analítica y un dinamizador de código para optimizar tiempos de despliegue.

IoT Cloud Einstein Plataforma de código cerrado de propiedad de la empresa Salesforce, cuyo fuerte es la inteligencia artificial para extraer información de los datos almacenados en su plataforma. Permiten ingresar datos provenientes de cualquier fuente en formato RESTful.

Watson IoT Es la plataforma IoT perteneciente a la empresa IBM, su fuerte es la marcada integración con las herramientas de analítica de IBM Blumix la cual permite realizar análisis de sentimientos, predicción y unir los datos con un grupo completo de herramientas de inteligencia artificial para obtener conocimiento del proceso.

Telit IoT Platform Plataforma IoT de la empresa Telit, su fuerte es la seguridad de los datos, contando con protocolos de encriptamiento punto a punto. Permite manejo remoto de los dispositivos sensados creando lazos de control para los mismos.

Azure IoT Plataforma IoT de Microsoft la cual se integra con los servicios Cloud de la misma empresa. Su solución de analítica es georeferenciada lo que le da un importante valor en la descripción de cambios geográficos, también resalta la facilidad de uso, generando configuraciones de todo el sistema en minutos.

Xively IoT Es la plataforma IoT perteneciente a la empresa Xively, realizan acompañamiento a las empresas para el levantamiento de hardware necesario para interconectar las variables físicas de interés a la plataforma y adicional a las características generales de las diferentes plataformas, esta ofrece conectividad con el CRM empresarial, lo que brinda integración con los procesos de logística y compras de la empresa.

Ayla IoT Es la plataforma IoT perteneciente a la empresa Ayla Networks, cuenta con asistencia a implementación de hardware, nube privada para el manejo de los datos de sus usuarios y aplicación móvil para el manejo remoto de dispositivos.

Oracle Internet of Things Plataforma IoT de la empresa Oracle, tiene las características básicas atribuibles a un sistema IoT con un punto adicional, la integración de Big Data a los desarrollos producto del liderazgo de Oracle en el manejo de bases de datos.

Fusion Connect Es la plataforma IoT de la empresa Autodesk, ofreciendo un Middleware para la integración de dispositivos de hardware además de un plugin para el desarrollo de inteligencia artificial sobre los datos registrados.

Predix Es la plataforma IoT de la empresa General Electric enfocada en el internet de las cosas empresarial combinando big data, predictiva, analítica y fácil desarrollo de aplicaciones.

AWS IoT Plataforma IoT perteneciente a la empresa Amazon, una de las más completas debido a la integración con técnicas de servicios flexibles en la nube. Cuentan con un core para instalar en más de 100 tipos de hardware y brindar comunicación con su nube integrando servicios de la aplicación Lambda para integración IoT, ofrecen un sistema operativo propio FreeRTOS para sus despliegues, además de esta integración con las herramientas de desarrollo de la suite Amazon.

WSO2 Plataforma IoT de código libre, tiene un core que permite conectar dispositivos de hardware destacando integración con Android, IOS, Windows, Arduino, Raspberry bajo licencia Apache. Soporta protocolos como MQTT, HTTP, Websockets y XMPP y adiciona módulos de analítica al sistema.

KaaloT Plataforma IoT de código abierto con las funciones generales de una plataforma IoT, la robustez de la misma radica en su comunidad de desarrollo y en su exhaustiva documentación. Integrable con cientos de dispositivos incluso con menos de

10K de RAM, disponible para trabajo en Java, C++ y C integrable con protocolos como Bluetooth, Zigbee o Z-Wave.

Thingsboard Plataforma IoT de código abierto cuyo fuerte es la telemetría y la integración de mapas a los cuadros de mando de la aplicación.

Macchina.io Esta plataforma IoT proporciona un entorno de tiempo de ejecución de JavaScript y C++ habilitado para web, modular y extensible. También es compatible con una amplia variedad de sensores y tecnologías de conexión, incluyendo Tinkerforge, Bricklets, Xbee y muchos otros, incluidos los acelerómetros.

Zetta Zetta es una plataforma orientada a servidores que se ha desarrollado en torno a Node.js, REST y una filosofía de desarrollo de programación reactiva basada en el flujo vinculada con las API de hipermedia de Siren. Están conectados con servicios en la nube después de ser agrupados como API REST. Incluye herramientas de visualización y soporte para análisis de máquinas como splunk.

ThingSpeak Plataforma de IoT que le permite analizar y visualizar los datos en MATLAB y elimina la necesidad de comprar una licencia. Recopila y almacena datos de sensores en canales privados y permite compartirlos en canales públicos.

DeviceHive Plataforma IoT de código abierto distribuida actualmente bajo la licencia Apache 2.0. Proporciona opciones de implementación en Docker y Kubernetes; se puede descargar y usar con nube pública o privada. Permite ejecutar análisis por lotes y aprendizaje automático, las bibliotecas de Android e iOS son compatibles en DeviceHive.

COMPARATIVA DE PLATAFORMAS IoT

El proyecto tiene limitantes geográficas debido a su instalación en zonas aisladas del territorio rural donde no se presentan sistemas de conectividad adecuados. La propuesta para solucionar este percance es utilizar una alternativa de código abierto que pueda ser instalada de forma local sin necesidad de salir a la nube, creando un ecosistema de dispositivos locales según la propuesta de la computación de techo. Por tanto se realiza la comparativa de las plataformas IoT de código abierto en los cuales es posible modificar su arquitectura como criterio del desarrollo del proyecto, con sistemas de desarrollo in situ en la tabla 2.2.

2.5.6. ESTÁNDAR W3C PARA SSN

Las ontologías son definiciones formales de tipos, propiedades y relaciones entre entidades; catalogan las variables requeridas para un conjunto y establecen relaciones entre ellas permitiendo limitar la complejidad y organizar la información.

La W3C Semantic Sensor Network Incubator Group (W3C) desarrollo un estándar de ontología para Semantic Sensor Network (SSN) basado en la necesidad de realizar consultas de información en redes de sensores. Tomo como base cuatro lenguajes para la descripción de sensores que presentaban la información en XML (W3C-Incubator-Group, 2011).

Como se mencionó anteriormente la base del estándar son lenguajes de descripción de sensores, entre ellos:

Plataforma	Control de dispositivos	Integración	Seguridad	Protocolos para colección de datos	Analítica	Soporta visualización
KaaloT	Si	SDK disponible portable para cualquier plataforma, REST API	SSL, llave RSA 2048 bits, llave AES de 256 bits.	MQTT, CoAP, XMPP, TCP, HTTP	Analítica en tiempo real, Cassandra BD	Si
Thingsboard	Si	REST API	SSL	MQTT, CoAP y HTTP	Apache, Spark, Kafka	No
WSO2	Si	REST API	SSL	MQTT, HTTP, ESB, WSO2	Servidor de analítica WSO2	Si

Cuadro 2.2: Comparativa de plataformas IoT de código abierto

SML Sensor Model Language. Basado en XML.

OM Observations and Measurements. Basado en XML, presenta fortalezas en guardado de información y manejo de datos casi en tiempo real.

TML Transducer Markup Language. Basado en XML.

Entre los servicios que se heredaron de estos lenguajes se tienen:

SOS Sensor Observation Services

SPS Sensor Planning Services

SAS Sensor Alert Services

El desarrollo del estándar abarco cuatro grandes ecosistemas tecnológicos, los dispositivos de sensado, las API's y redes de sensores, los servicios de aplicación y los servicios Mashup. Los dos primeros no tienen un estándar propietario sino son de libre creación por los desarrolladores, los servicios de aplicación corresponden a la Sensor Web y son desarrollos enfocados en datos por la OGC y los servicios de Mashup corresponden a la Web Semántica y son desarrollados por la W3C. La forma en que interactúan en el desarrollo del estándar son presentados en la figura 2.22.

En la etapa de descubrir datos y enlazarlos los que se propone es la adecuación de los datos a la ontología para luego asociar estos datos a unos dispositivos que brindan

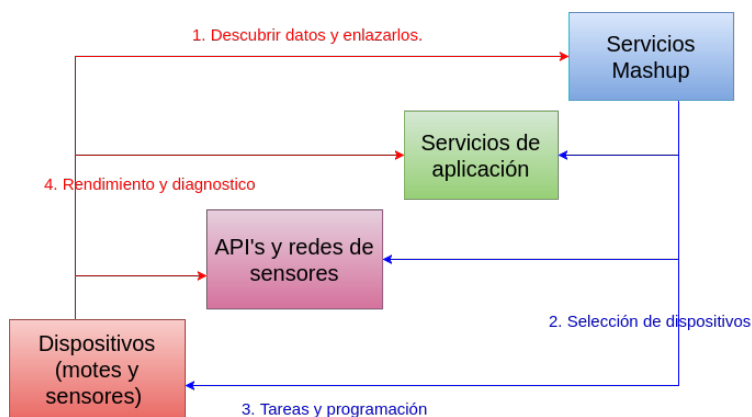


Figura 2.22: Tecnologías y casos de uso
Fuente: Adaptado de (W3C-Incubator-Group, 2011)

la infraestructura necesaria para levantar los datos y crear de estos tareas y ejecutar programación de estas o de procesos asociados. Todo este proceso se realiza para la jerarquización de los datos y que sean consumidos por un Mashup que es un sistema que consume datos de diferentes fuentes con un fin en específico.

Para la creación del estándar se revisaron 17 ontologías basadas en sensores, entre ellas:

- CSIRO Sensor Ontology
- Onto Sensor
- Swamo
- SDO
- MMI Device Ontology
- Sensor ML Processes
- CESN
- WISNO
- Ontonym-Sensor
- Seek OBOE
- Seres O&M
- Stimuli-Confered
- Sensei Observation and Measurement Ontology
- OOS Thethys

- O&M- OWL (SemSOS)
- Socio-Ecological Research and Observation oNTology SERONTO

ONTOLOGÍA SSN

La ontología desarrollada se basa en los conceptos de sistema, proceso y observaciones. Es de importancia remarcar que los sensores no están limitados a dispositivos de medidas físicas, ya que por definición un sensor es algo que puede estimar o calcular el valor de un fenómeno, así un dispositivo o un proceso computacional puede ser incluido en esta descripción.

La estructura de la ontología se agrupa en módulos, clases y relaciones, los módulos se presentan en la figura 2.23, las clases y relaciones en la figura 2.24.

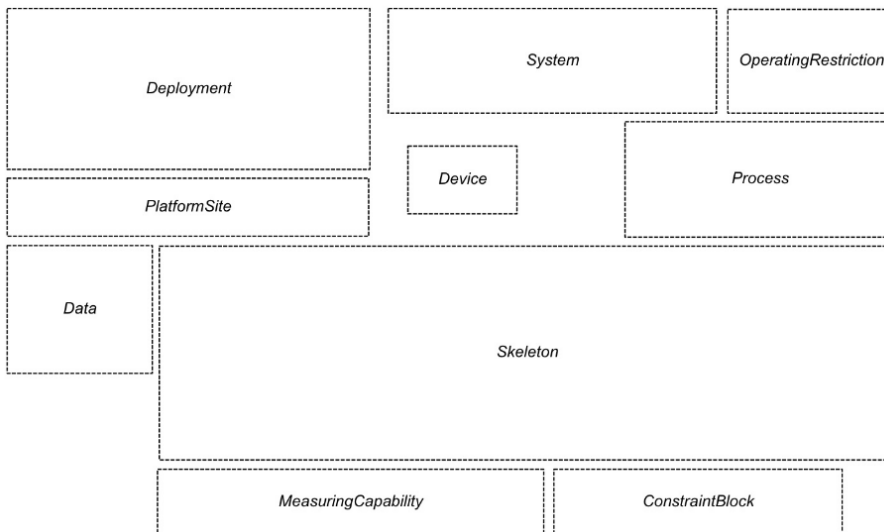


Figura 2.23: Módulos de la Ontología SSN
Fuente: (W3C-Incubator-Group, 2011)

Esta ontología puede ser usada desde diferentes perspectivas.

- Perspectiva del sensor que indica como se sensa y que es sensado.
- Perspectiva del dato, indica las observaciones realizadas y los metadatos del sistema.
- Perspectiva del sistema, se enfoca en como se relacionan los sistemas de sensores.
- Perspectiva de características y propiedades, que busca especificar que se puede sensar a través de las propiedades y características del arreglo.

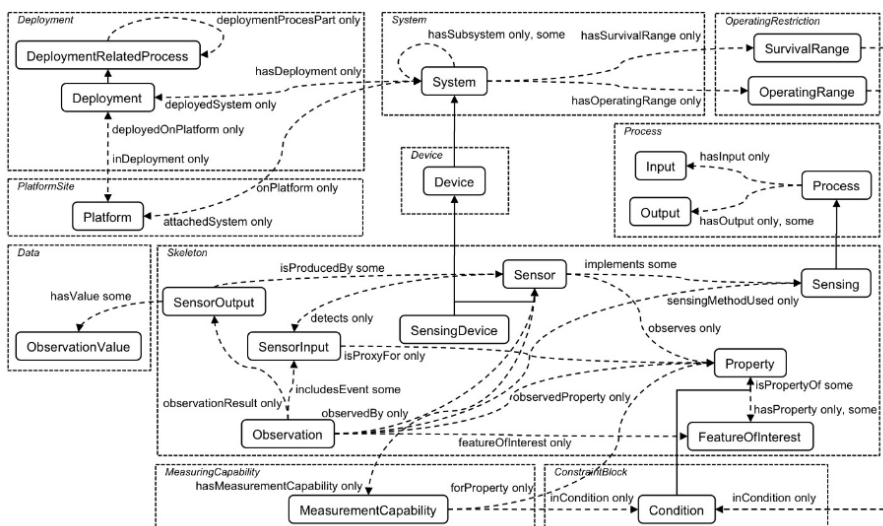


Figura 2.24: Vista de la Ontología SSN en clases y propiedades
Fuente: (W3C-Incubator-Group, 2011)

2.6. ESTUDIOS SIMILARES

El análisis y estudio de formas para optimizar el trabajo en campo por medio de tecnología a sido ampliamente desarrollado y sus experiencias consignadas en fuentes de difusión científica. En esta sección se consignan los trabajos desarrollados en áreas como el despliegue de redes de sensores inalámbricas en espacios rurales, la integración de estas redes con imágenes provenientes de drones o diferentes formas de aproximación al problema de riego inteligente. También existen aplicaciones de IoT en el campo de monitoreo de sistemas agrícolas.

2.6.1. SENSADO DE PARÁMETROS RELEVANTES EN AGRICULTURA

La forma de capturar datos al momento de realizar procesos agrícolas es un campo de estudio pues la forma en que los sensores interactúan, son localizados o como influyen sus sensibilidades son relevantes para el desarrollo de un sistema robusto.

La conexión inalámbrica presente en la etapa de sensado a sido discutida por varios autores en donde han definido las Red de Sensores Inalámbricos (Wireless Sensor Networks) (WSNs) híbridas como promisorias dentro del sensado de variables importantes en granjas. La diferencia entre un sistema híbrido y uno terrestre es la red inalámbrica de sensores subterráneos que se comunican en el suelo, esto buscando eliminar una etapa de comunicación para sensores de suelo que salgan a la superficie pudiendo interferir en las acciones del agricultor y el desarrollo del proceso agrícola. Se desarrollo un marco de trabajo sobre tierra formada de un 50 % arena, 35 % limo y 15 % arcilla. El experimento relaciono varios tipos de densidad de la tierra desde 5 % hasta 25 % y tres frecuencias diferentes de transmisión (433, 868 y 915 MHz) (Yu et al., 2013) teniendo en cuenta las pérdidas de canal, la tasa de error de bits y la distancia máxima de transmisión

en variables como tipo de tierra, contenido volumétrico de agua y distancia entre nodos, se mostró que las pérdidas de la señal de radio es menor en bajas frecuencias y baja densidad en la tierra, además el despliegue de un nodo profundo afecta la atenuación de la señal a 433 MHz determinando la mejor manera de realizar un nodo de comunicaciones a diferentes profundidades.

Esta conexión inalámbrica también es analizada desde la forma en que interactúa con la nube, la alta capacidad de generar datos junto al almacenamiento y procesamiento además de la elasticidad y el pago por uso hacen de la nube una solución viable a las implementaciones basadas en IoT para procesos agrícolas (Sales et al., 2015). Además de la nube se han realizado implementaciones en servidores locales buscando de emular el concepto de computación en la niebla o en el techo mostrando que las aplicaciones web de bases de datos y sistemas móviles facilitan la adquisición en tiempo real para un efectivo monitoreo.

Estas herramientas cuando son implementadas con soluciones de código abierto ahorran dinero y permiten un alto grado de integración y de robustez. Juntando aplicaciones Android con sistemas de control almacenados en forma local por medio de un servidor basado en Linux, Apache, MySQL, PHP, Perl o Python se tiene una solución para el trabajo con grandes volúmenes de datos (Montoya et al., 2013), los nodos sensores se adaptan con una versión de sistema operativo para pequeños dispositivos conocido como TinyOS y montado sobre un sistema de comunicaciones como es 6LoWPAN dando resultados en los cuales la arquitectura se muestra eficiente para el análisis de los datos levantados. De una forma similar en España se ha propuesto la necesidad de instrumentar variables tales como humedad y nivel de agua para adaptar técnicas asociadas de control buscando optimizar la producción agrícola e implementando la comunicación sobre GPRS y la fuente de alimentación paneles solares garantizando su uso en zonas que no cuentan con un eficiente fluido eléctrico (Navarro-Hellin et al., 2014), mostrando una tasa de consumo por sensor de 6 mA y un total de 300 mA, con un consumo de 1 A cuando se entabla conexión con la estación base; sin paneles solares la solución podría operar durante 13 días seguidos. Resultados similares se constataron en Algeria pero sin implementación, validados por software (Khelifa et al., 2015).

Un ejemplo adicional es propuesto desde el departamento de Ingeniería Eléctrica de la Universidad de Norte de Texas en donde diseñaron una WSN usando Arduino como tarjeta de sensores y Raspberry Pi como estación base (Ferdoush and Li, 2014). La estación base montada en la Raspberry se comunica por Zigbee con los sensores en donde los datos llegan en primera instancia al gateway para ser transportador a un servidor de bases de datos. Este a su vez alimenta una aplicación web realizada en PHP sobre el servidor web y presenta los resultados en una interfaz HTML que puede ser accedida por Ethernet o Wifi a través de un router.

2.6.2. PROPUESTAS DE ARQUITECTURA

El sistema de sensado de datos meteorológicos es una de las formas de estimar la evapotranspiración de un cultivo y así proyectar sus necesidades hídricas, proceso explicado previamente. Los actuales sistemas inalámbricos de sensado no tienen el sistema de recolección de datos, acceso ni la forma en que se comparten estandarizados, lo que impide el uso de datos provenientes de diferentes redes. Sawant propone la incor-

poración de un marco de trabajo basado en los estándares del Consorcio Geoespacial Abierto (OGC - Open Geospatial Consortium) para el manejo web de sensores. La propuesta incluye hardware y software en donde diseña una plataforma de hardware donde integra medidas agro-meteorológicas, una interfaz web y estima la evapotranspiración del terreno por medio de la ecuación Penman-Monteith (Suryakant et al., 2017).

Existen algunos intentos en definir el estado actual de integración entre el paradigma IoT con campos como la agricultura. Algunos autores han explorado este concepto mostrando alcances y necesidades específicas del sector como Pratim que revisa varias aplicaciones del IoT y los retos asociados con el desarrollo de tecnologías para mejorar la granja teniendo como base que estos desarrollos deben ser baratos para poder ser apropiados por el grueso de los agricultores. Propone un esquema para agricultura inteligente basado en seis capas, una capa física concerniente a sensores y levantamiento de datos, una capa de red que abarca los protocolos necesarios para comunicar estos datos una capa intermedia o Middleware que se refiere a una pasarela en donde los datos son recibidos y tratados para extraer información de ellos, una capa de servicios en donde se analiza la información en tiempo real para tomar decisiones de campo, una capa analítica encargada de hacer análisis predictivo para descubrir futuros desafíos y generar análisis multivariado y una capa de experiencia de usuario (Ray, 2017), esta arquitectura se presenta en la figura 2.25.

Esta propuesta se plantea para usos como irrigación, control de plagas y enfermedades, monitoreo de movimiento del ganado, monitoreo diario, registro de calidad de agua, seguimiento de gases, de condiciones del suelo, de agricultura de precisión por medio de drones. También relaciona los retos futuros a solucionar y la hoja de ruta para realizarlo; entre los retos ubica el costo del desarrollo, la estandarización, heterogeneidad, el desarrollo de un contexto, la identidad de nodo, el manejo de la energía, la tolerancia a fallas y la necesidad de ser una solución en tiempo real.

Surgen propuestas de controladores on-off para el control de válvulas en donde se diseño el firmware y actuador el cual a través de un gateway realiza un lazo de control a 54 válvulas ubicadas a lo largo de 1610 metros alimentado por un panel solar; se le realizaron al sistema pruebas de estrés donde se determino una vida útil de las válvulas de cuatro años con una recuperación de la inversión de tres años (Coates et al., 2013). Un desarrollo similar se encuentra en cultivos de naranja en Chongqing, China (Zhang et al., 2017).

En Galicia fue presentado en el Congreso nacional del Medio Ambiente el proyecto SmartCrops, el cual se basa en construir una solución de telemetría basado en sensores con tecnología inalámbrica como herramienta facilitadora de técnicas de agricultura de precisión incorporando conectividad IP y tecnologías Web y GIS y minería de datos para analizar la información recogida por la aplicación (Iglesias et al., 2012).

Las arquitecturas propuestas para agricultura de precisión barajan diversos tipos de variables a ser sensadas, Prakash indica que las recomendaciones agrícolas se basan actualmente en la acumulación de datos, aunque es la realidad sea un proceso lento y en Australia levantado de forma manual. Presenta una plataforma IoT basada en el proyecto de código abierto OpenIoT proyectada para levantar información ambiental, fertilización e irrigación teniendo la opción de correlacionar estos datos y extraer información del mismo. Los interesados en esta plataforma además de ser los agricultores son bo-

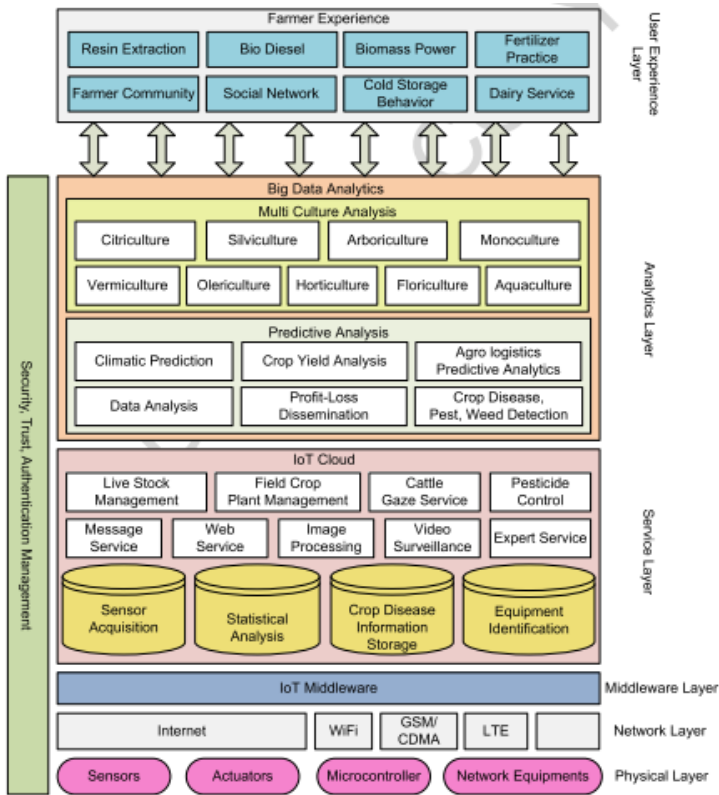


Figura 2.25: Arquitectura de manejo de granja bajo el paradigma IoT
Fuente: (Ray, 2017)

tánicos, biólogos, gobierno y organizaciones comerciales que soportan operaciones de granja pues adicionalmente puede integrar medidas de pH o cantidad de nitrógeno en ciertas condiciones (Jayaraman et al., 2016).

Esta plataforma de nombre SmartFarmNet provee un laboratorio virtual medioambiental que maneja grandes volúmenes de datos soportando consultas rápidas, interacción del usuario en tiempo real y la habilidad de compartir datos de estudios hechos por múltiples investigadores. Los sensores usados se adicionan a la plataforma siguiendo los lineamientos ontológicos de la *Semantic Sensor Network SSN* la cual provee un esquema que integra sensores, observaciones, atributos y otros conceptos relacionados. La información la almacena en una base de datos NoSQL sobre una plataforma elástica. La arquitectura de la herramienta se presenta en la figura 2.26.

Se aprecia la arquitectura de la solución SmartFarmNet desarrollada por los autores con el módulo de creación de consultas para generar información, el marco de trabajo para la descripción del recurso y la interfaz de programación.

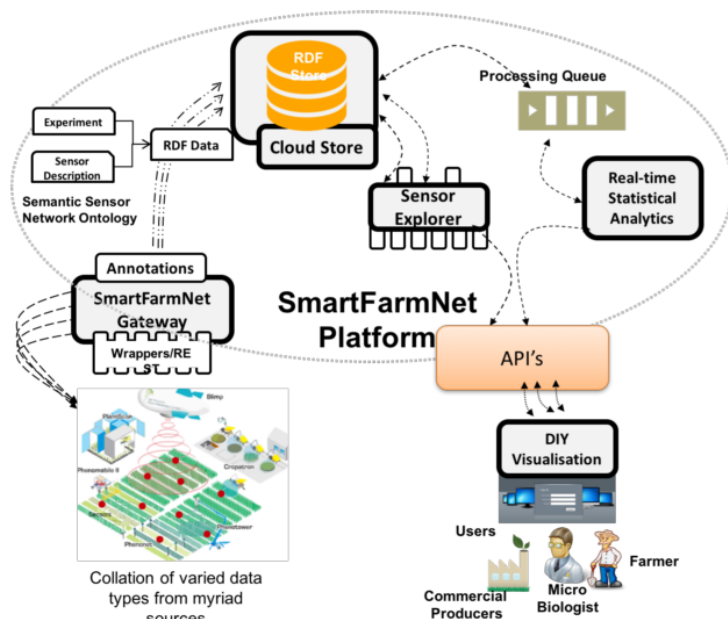


Figura 2.26: Arquitectura de SmartFarmNet
Fuente:(Navarro-Hellin et al., 2016)

2.6.3. ESTIMACIÓN DE RIEGO

El IoT a sido usado en conjunto con diferentes herramientas para solucionar el problema de riego inteligente. Cambra y Sendra proponen un sistema llamado PLATEM PA el cual es alimentado por medio de sensores e imágenes provenientes de drones con cámaras multispectrales las cuales generan índices NDVI y por medio de una arquitectura orientada a servicio estiman las necesidades hídricas de un campo, el método que usan es el balanceo de energía para saber la cantidad de agua que requiere el campo, así que los sensores están enfocados en levantar información en su mayoría de las condiciones del suelo (Cambra et al., 2017).

Existen propuestas para la medición de humedad del suelo en forma descentralizada (Smarsly, 2013) en donde por medio de nodos de sensores de bajo costo y agentes de software móviles se analizan parámetros de humedad para unirlos con variables climáticas para predecir disparos de eventos de irrigación. Los nodos están basados en la arquitectura de Oracle SunSPOT y desarrollado sobre Java.

En otras latitudes la estimación de riego se realiza solamente con las ecuaciones básicas descritas por (FAO, 2017) pero acompañadas de una infraestructura tecnológica que la soporte (Rajalakshmi and Mahalakshmi, 2016). Rajalakshmi y Mahalakshmi utilizan Arduino para recibir datos de humedad del suelo, temperatura y luminosidad para ser enviados a un servidor web por medios inalámbricos en formato JSON y datos de un sensor ultrasónico para registrar el nivel del tanque de agua de almacenamiento disponible para el cultivo, además desarrollaron una aplicación sobre Android que accede a

esta información y permite controlar de forma manual una bomba de riego y enviar notificaciones al agricultor sobre el estado del cultivo. Para la transmisión usa un módulo Ethernet, los servicios web fueron realizados en PHP. El control usado para sistematizar el riego es un controlador on-off, el cual activa la bomba de riego cuando la humedad de la tierra baja de un umbral preestablecido.

Otra alternativa de estimación de riego a sido propuesta por Kavilan el cual levanta por medio de una WSN parámetros de tipo de suelo, contenido, temperatura y humedad. Registra el color de las hojas de las plantas y luego usa el análisis de discriminante lineal de Fisher para determinar cuando la planta se encuentra en óptimas condiciones o en su caso contrario necesita ser aplicada agua al cultivo. Para la parte final y verificar la información levantada utiliza una máquina de vector transductiva para cuantizar los cambios entre las medidas y no generar riegos innecesarios (Kabilan and Selvi, 2016).

En la King Mongkut's University of Technology Thonburi (KMUTT) proponen un sistema de granja inteligente enfocado en mejorar el proceso de siembra compuesto de dos partes, una de sensado y una de control. La parte de control se centra en riego en un sistema bajo techo basado en datos estadísticos de sensores de temperatura, luminosidad y humedad; para que los datos no tengan ruido se aplica un filtro Kalman para suavizarlos y luego ser introducidos al sistema de control, el cual es un árbol de decisión basado en una serie de reglas asociadas a los sensores el cual compara la información producida por estos con unos umbrales de decisión para accionar una válvula (Putjaika et al., 2016), generando un controlador de múltiples entradas y única salida (MISO).

El trabajo mas avanzado en el campo de uso de inteligencia computacional enfocado al riego es el propuesto en España por investigadores de la universidad de Cartagena, la universidad de Belfast en UK y empresas que venden soluciones de tecnología como Widhoc Smart Solutions en Murcia las cuales desarrollaron un sistema de soporte de decisiones en riego inteligente de frecuencia semanal recogiendo información de sensores ubicados en diferentes partes del cultivo cruzándola con datos meteorológicos formando un lazo de control cerrado adaptado a soportar perturbaciones y corregir errores, para hacer esto propone el uso de dos técnicas de inteligencia computacional, las PSLR (Partial Least Square Regression o Regresión Parcial por Mínimos Cuadrados) y las ANFIS (Adaptive Neuro Fuzzy Inference Systems o Sistema de Inferencia Adaptativa Neuro-Difusa) (Navarro-Hellin et al., 2016).

Postulan que la formulación propuesta para estimar la evapotranspiración por la FAO es un sistema de control en lazo abierto (FAO, 2017) lo que significa que el modelo es susceptible a ser modificado por eventos estocásticos propios del entorno y no es propenso a corregir las perturbaciones locales; para solventar esto comúnmente se requiere un agrónomo experto que interprete esos cambios y corrija el modelo debido a la complejidad del mismo.

Esta dinámica lleva a proponer un reemplazo al factor humano en cuanto a la interpretación de estos factores propios locales por medio de sensores en campo que los registren con una tasa de muestreo de 15 minutos y un sistema experto basado en inteligencia computacional para solventar el proceso de interpretación de la información y toma de decisiones. El diagrama de como se realiza el proceso sin intervención de un sistema experto es presentado en la figura 2.27.

Esta gráfica muestra como el agrónomo recibe información del clima, las caracterís-

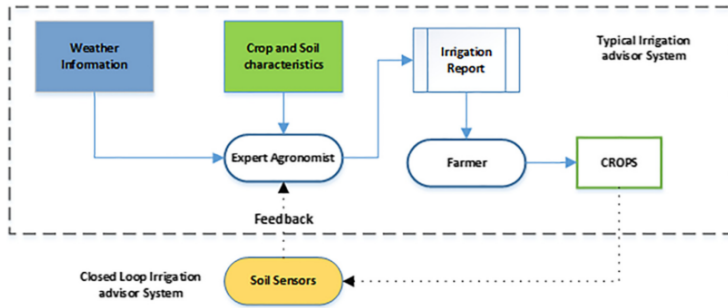


Figura 2.27: Sistema de cálculo para necesidades de riego

Fuente: (Navarro-Hellin et al., 2016)

Las características propias del terreno y los sensores en tierra para generar un reporte de necesidades de irrigación, el cual es suministrado al agricultor y este aplica las recomendaciones a su cultivo en un lazo cerrado de control.

3

DATOS Y MÉTODOS

3.1. INTRODUCCIÓN

Todos los desarrollos, códigos y material generado se encuentra en el repositorio en el portal web <http://ritaportal.udistrital.edu.co:10257/> en donde está alojado el sistema de estimación de riego.

El desarrollo de este proyecto se fundamentó en el método de investigación científica descrita por Roberto Hernández Sampiere en su obra “Metodología de la investigación” (Hernández Sampieri et al., 2003) el cual describe una serie de pasos para el desarrollo de investigaciones a la cual se le añadió una serie de pasos intermedios para la validación parcial de los resultados y el cumplimiento de los objetivos.

La metodología desarrollada inicia con el diseño de la investigación que es el método por el cual se logró la abstracción del problema para continuar con la descripción de las variables a ser utilizadas y continuar con la explicación de la población en la cual realizar el análisis.

La metodología desarrollada define la muestra en la cual realizar la validación de la solución y describe los instrumentos y procedimientos que serán usados para desplegarla. Se finaliza con la documentación técnica del resultado el cual satisface uno de los objetivos específicos propuestos.

Esta es la manera en que el proyecto fue ejecutado y cada etapa se pasa a describir a continuación.

3.2. DISEÑO DE LA INVESTIGACIÓN

Ya definidos los conceptos y alcance en el planteamiento del problema y marco teórico desarrollado se abordó el diseño de la investigación por el método de análisis y síntesis basado en los desarrollos propuestos por (Suryakant et al., 2017), (Ray, 2017) y (Navarro-Hellin et al., 2016) donde los autores crean un flujo de información constante entre las variables físicas registradas y el sistema propuesto que extrae datos relevantes de las diferentes fuentes consultadas. A los modelos propuestos por los autores se agregaron variables meteorológicas para crear un lazo cerrado de información y calcular la

evapotranspiración de referencia además se usaron bases de datos con información medioambiental relevante y se documentó el proceso de incorporarla al sistema de gestión de conocimiento. Esta base de datos se complementó con un modelo de inteligencia artificial para la correcta estimación de la evapotranspiración y poder calcular las necesidades hídricas del cultivo al cual se fuera a hacer referencia.

3.3. VARIABLES

Las variables son conceptos que asumen diferentes valores y se refieren a aquellas propiedades, atributos, características, magnitudes, funcionalidades, etc, que podrían estar presentes en el universo de estudio. Debido al problema que se abordó las variables son de naturaleza cuantitativa y según su función son independientes, cuyo fin en el proyecto fue la creación de variables dependientes para la estimación de riego inteligente. Las variables que se utilizaron son simples en su gran mayoría pues solo fue necesario un indicador para expresar la magnitud a la que hacían referencia excepto las vinculadas al proceso de inteligencia computacional que por su naturaleza requirieron más de un indicador para describirlas, esto las cataloga como variables complejas.

El proceso de descripción de las variables se ajustó a los lineamientos desarrollados por el W3C, grupo de trabajo creado para realizar la descripción de ontologías para sensores y redes de sensores (Barnaghi et al., 2011).

3.4. POBLACIÓN Y MUESTRA

El proyecto es concebido como una solución general cuya población objetivo son los agricultores, por lo tanto su área de aplicación es global a pesar de haber sido enfocado en una problemática Colombiana. Para el desarrollo de los experimentos se tomó como muestra una comunidad ubicada en Caldas de nombre La Dorada que contó con apoyo estatal para el levantamiento de información medioambiental debido a su condición de población vulnerable víctimas del conflicto armado. Un desarrollo mayor de la población y las condiciones tanto sociales como técnicas del sector se realiza en el desarrollo del proyecto.

Esta se ubica en la zona espacial correspondiente a latitud 5.4599237 y longitud -74.66312408.

3.5. INSTRUMENTOS Y PROCEDIMIENTOS

Para el desarrollo del proyecto se realizó una búsqueda de datos medioambientales en fuentes nacionales tales como empresas del sector que tienen visores para los datos que levantan en campo, organismos gubernamentales de monitoreo y control de fenómenos naturales, organizaciones no gubernamentales que disponen de fuentes de datos afines y por último organizaciones internacionales que proveen información similar. El levantamiento de estas fuentes de datos se realizó por medio de la observación partiendo del Sistema de Información Ambiental de Colombia (SIAC) y realizando un reconocimiento de los asociados a la misma, luego se partió de empresas que ofrecen estaciones meteorológicas para determinar cuáles tienen una interfaz de datos abiertas, cuáles no y las que lo tienen abierto como es la estructura de los datos para integrarlos al modelo

de gestión del conocimiento. Para la descripción de los datos se utilizó la propuesta por el W3C.

Para determinar la plataforma se priorizó el poder adaptarla a las necesidades del proyecto y que fuera agnóstica con respecto al hardware recibiendo cualquier tipo de dato proveniente de diferentes servicios, esto llevó a la utilización de una plataforma meteorológica proveída por la empresa DAVIS que presenta los resultados de sus mediciones bajo el formato abierto JSON, brindando una interfaz web que permite la lectura de la estación en tiempo real, lo que permite guardar los reportes y registros de la información para su posterior análisis.

El modelo de gestión de conocimiento usado parte de priorizar procesos para la descripción de datos tecnológicos, la forma en que interactúan y la forma en generar conocimiento y sabiduría de los mismos. Se modificó el mismo para adecuarlo a las necesidades propias del proyecto.

El sistema de inteligencia computacional usado se escogió basado en la forma de los datos conseguidos partiendo del criterio de sistema de aprendizaje supervisado o no supervisado. Los datos por su estructura permitieron alimentar un modelo para la generación de conocimiento del sistema.

3.6. ANÁLISIS DE LOS DATOS

Los datos como recurso básico del sistema de gestión del conocimiento planteado son fundamentales para el propósito de la investigación y al ser de naturaleza cuantitativa su análisis parte de comparaciones numéricas y estadísticas para discernir los rangos de trabajo del sistema. Estos se ubican por categorías dándole un orden debido a su importancia o unidad de medida.

Luego de que los datos son convertidos en información se relacionan con el sistema de gestión y las reglas de inteligencia computacional las cuales generan estimaciones de riego, siendo la salida del sistema un valor numérico y una orden de activación del sistema de riego.

3.6.1. SENSORES

Las diversas soluciones en el mercado para realizar mediciones de las variables presentadas abarcan desde el movimiento DIY (Hazlo Tu Mismo) basados en Arduino, Raspberry u otros microcomputadores disponibles hasta soluciones robustas de hardware como estaciones meteorológicas que integran los sensores en un solo dispositivo para realizar medición climática y meteorológica. Dentro de los instrumentos tenidos en cuenta para recoger los datos en campo se priorizó los que tienen certificación vigente en el país por el Organismo Nacional de Acreditación de Colombia (ONAC) que es el instituto encargado de certificar los sensores utilizados por los instrumentos de medición para validar que sus medidas sean idóneas para el proceso referenciado, esta responsabilidad la realizan con el Instituto Nacional de Metrología (INM) y se consideró de relevancia pues el sistema es sensible a cambios si las medidas no corresponden a la realidad del conjunto (de Acreditación Colombiana, 2007).

Se encontró que ninguno de los sensores presentes en los desarrollos investigativos eran validados por la ONAC al igual que las empresas que ofrecen mediciones meteoro-

lógicas y cuentan con un hardware propio lo tienen validado ante el estado.

Con base en esto se optó por una estación meteorológica validada en el extranjero con las normas de sensores respectivas y que tuviera presencia en el país. Las normas son las NIST que hacen referencia a el National Institute of Standards and Technology de EEUU, adicional a esto existe una empresa con presencia en Colombia que distribuye sensores y estaciones certificadas, ello llevó a la escogencia de la estación meteorológica MA3081 PRO de la empresa Meteogadro desarrollada por Davis con los sensores certificados presentados en la tabla 3.1.

3

SENSOR	PRECISIÓN	SENSOR COMPROBADO Y PRECISIÓN VERIFICADA	PUNTOS DE VERIFICACIÓN
Humedad interna	+- 0.1 RH a 33% RH, +- 0.3% entre 80% RH y 90% RH	+- 3% RH entre 33% y 90%	33% RH, 80% RH y 90% RH
Humedad externa			
Temperatura interna	+- 0.4 °F (0,2°C)	Precisión en función de la temperatura	-40°F, -15°F, 0°F, 40°F, 80°F y 140°F (-40°C, -26°C, -18°C, 4°C, 27°C y 60°C)
Temperatura externa		Precisión en función de la temperatura	
Presión barométrica	+- 0.03 pul Hg (0.1 hPa)	± 0.03 pul Hg (1,0 hPa) entre 20 pul Hg y 30 pul Hg (677 hPa y 1016 hPa)	20.00 pul Hg, 22.25 pul Hg, 25.00 Hg, 27.75 Hg y 30.00 pul Hg) (677 hPa, 753 hPa, 847 hPa, 940 hPa y 1016 hPa
Velocidad del viento	± 2 mph (0,9 m/s)	± 2 mph (0,9 m/s) sobre 40 mph (18,0 m/s) ± 5% sobre 40 mph (18,0 m/s)	6, 12, 25, 40 y 75 mph (2,7; 5,4; 11,2; 18,0 y 33,5 m/s)
Lluvia 0.01 pul	± 0,07 ml = ± 0.001" (0,003 mm) lluvia = 1/100 de boquilla	± 4% + 1 boquilla (arriba de 4"/hr)	1.00 pul total rain at rate of 2.21pul/hr
Lluvia 0.2 mm		± 4% + 1 tip (tasa de lluvia 100 mm/hr)	25,4 mm total de lluvia a una tasa de 56 mm/hr

Cuadro 3.1: Sensores certificados por la empresa Davis en EEUU

FRECUENCIA DE LOS DATOS

La frecuencia refiere a la cantidad de tiempo que pasa entre una medición y otra, el sistema propuesto cuenta con las presentadas en la tabla 3.2.

SENSOR DE TEMPERATURA Y HUMEDAD

Los sensores utilizados en la medición de temperatura y humedad con sus características técnicas se aprecia en la tabla 3.3.

DATO	FRECUENCIA
Velocidad del viento	2,5 a 3 segundos
Dirección del viento	2,5 a 3 segundos
Lluvia acumulada	20 a 24 segundos
Intensidad de lluvia	20 a 24 segundos
Temperatura exterior	10 a 12 segundos
Humedad exterior	50 segundos a 1 minuto
Radiación solar	50 segundos a 1 minutos

Cuadro 3.2: Frecuencia de datos de sensores

SENSOR DE RADIACIÓN SOLAR

Los sensores utilizados en la medición de radiación solar y sus derivados con características técnicas se aprecia en la tabla 3.4.

SENSOR DE VELOCIDAD DEL VIENTO (ANEMÓMETRO)

El sensor utilizado en la medición de la velocidad del viento y sus derivados con características técnicas se aprecia en la tabla 3.5.

Tiene dos variables principales, la dirección del viento que registra con un sensor tipo flecha el cual es acondicionado con un potenciómetro y la sensibilidad depende del potenciómetro integrado. La segunda variable es la velocidad del viento que registra con un sensor de copas con Switch magnético.

SENSOR DE HUMEDAD EN SUELO



El sensor utilizado en la medición de la velocidad del viento y sus derivados con características técnicas se aprecia en la tabla 3.6. Es un sensor de tipo resistivo que mide el contenido de agua en la tierra que tiene un rango de 0 (lo más seco) a 200 (lo más húmedo) centibars.

3.6.2. CAPTURA DE LA INFORMACIÓN DE LOS SENSORES

La información que entregan los sensores ubicados en territorio entregan información de manera constante en formato JSON, el cual es un formato abierto de notación en JavaScript de fácil lectura, siguiendo el estándar propuesto por la ontología se refieren las variables presentadas en la tabla 3.7.

Para la captura de la información se desarrollo un script en el lenguaje de programación R el cual desarrolla el siguiente algoritmo:

1. Se abre la conexión a la estación en internet .
2. Se descarga fichero de información.
3. Se cierra la conexión
4. Se verifica que la información queda guardada en la variable asignada.
5. Verificar integridad del dato en el JSON pues tiene diferentes dimensiones.

<p>El sensor de temperatura es tipo diodo de silicón:</p> <p>Rango -40 °C a 65°C</p> <p>Resolución 1°C</p> <p>Actualización 77 a 90 segundos</p>	
<p>Humedad relativa: Capacitor magnético tipo film, mide la humedad relativa de 1 a 100 %</p> <p>Rango 1 a 100%</p> <p>Resolución 1%</p> <p>Actualización 10 segundos</p>	

Cuadro 3.3: Sensores de temperatura y humedad

6. Extraemos la lista de utilidad que contiene los datos relevantes.
7. Convertir JSON en un data frame para operar en R como archivo salida del script
8. Agregar el último registro de la estación a el archivo salida generado
9. Restricción 1: Si es la primer vez que se ejecuta el script se debe crear previamente el archivo EstacionDataFrame compuesto por dos veces el último dato, esto para generar la estructura de la tabla de forma indirecta sin hacerlo con un motor de bases de datos.

Este almacena la información meteorológica sobre la cual se hace el análisis de cálculo de la evapotranspiración.

3.6.3. CALCULO DE LA EVAPOTRANSPIRACIÓN

El calculo de la evapotranspiración pasa por tres fases las cuales implican el cálculo por medios matemáticos derivados del conocimiento presentado previamente, la estimación del mismo por medio de un modelo de inteligencia computacional y el cálculo de la evapotranspiración ajustada para el tipo de cultivo que se quiera y la extensión del mismo. Las fases se desarrollan de la siguiente manera:

Este sensor capta parte del espectro UV y tiene una cubierta plástica ABS resistente UV.

Transductor Tipo fotodiodo

Respuesta 280 a 360 nm

Temperatura de operación -40°C a 65°C



Dosis de radiación:

Rango 0 a 199 MEDs

Resolución 0.1 MEDs a 19.9 MEDs

Actualización 10 segundos

Índice de radiación UV

Rango 0 a 16 del Índice

Resolución 0.1 del Índice

Exactitud +- 5% de la escala



Cuadro 3.4: Sensores radiación solar

Fase 1 Se realiza la medición de la evapotranspiración por medio de la ecuación de Hargreaves.

Fase 2 Al tener el resultado de salida los datos adquiridos son usados para encontrar el modelo de inteligencia computacional que mejor responda a la medida de evapotranspiración, realizado de esta forma se cuenta con un predictor por medio de inteligencia artificial.

Velocidad del viento:

Rango 1 a 322 km/h

Resolución 1 km/h

Exactitud +- 3 km/h

Dirección del viento:

Rango 0° a 360° o 16 puntos de compás

Resolución 1°

Exactitud +- 7

Wind Run:

Rango 0 a 2000 km

Resolución 0.1 km

Exactitud 0 +- 5



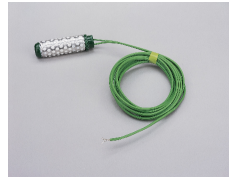
Cuadro 3.5: Sensor de velocidad del viento

Fase 3 El resultado provisto por el sistema de inteligencia artificial se utiliza para por medio de la constante de cultivo K_c calcular la evapotranspiración ajustada y con la magnitud del terreno estimar el consumo en litros de agua necesarios.

Resolución 1cb (centibar)

Rango 0 a 200 cb

Tiempo de actualización 77 a 90
segundos



Cuadro 3.6: Sensor de velocidad del viento

Variable	Significado
dewpoint_c	Temperatura de punto de rocío en grados Celsius
dewpoint_f	Temperatura de punto de rocío en grados Fahrenheit
heat_index_c	Índice de calor en grados Celsius
heat_index_f	Índice de calor en grados Fahrenheit
precip_rate_inh	Tasa de precipitación en pulgadas por hora
precip_rate_mmh	Tasa de precipitación milímetros por hora
precip_today_in	Lluvia acumulada en el día en pulgadas
precip_today_mm	Lluvia acumulada en el día en milímetros
pressure_hpa	Presión atmosférica en hectopascales
pressure_inhg	Presión atmosférica en milímetros de mercurio
pressure_mb	Presión atmosférica en milibares
relative_humidity	Humedad relativa del aire a dos metros sobre el suelo
report_date	Tiempo de generación de datos UTC
report_date_locale	Tiempo de generación de datos hora local
solar_radiation	Cantidad de radiación solar, vatios/metro cuadrado
temp_c	Temperatura del aire a 2 metros en grados Celsius
temp_f	Temperatura del aire a 2 metros en grados Fahrenheit
uv_index	Valor de radiación ultravioleta
wind_degrees	Dirección desde la cual proviene el viento en grados
wind_dir	Dirección desde la cual proviene el viento puntos cardinales
wind_gust_kmh	Máxima velocidad del viento en hm/h
wind_gust_mph	Máxima velocidad del viento en millas/h
wind_speed_kmh	Promedio de velocidad del viento en km/h
wind_speed_mph	Promedio de velocidad del viento en millas/h
windchill_c	Temperatura de enfriamiento del aire a 2 mt, grados Celsius
windchill_f	Temperatura de enfriamiento del aire a 2 mt, grados Fahrenheit
altitude	Altitud
brand	Marca de la estación o dispositivo que recoge datos
country	País
created	Fecha de creación de la plataforma en UTC
created_locale	Fecha de creación de la plataforma en hora local
elevation	Elevación
latitude	Latitud
location	Nombre de localización
longitude	Longitud
model	Modelo
state	Estado (En Colombia sería departamento)
station	Estación
timezone	Zona de tiempo local

Cuadro 3.7: Datos provenientes de sensores. Fuente: Propia.

4

DESARROLLO

Todos los desarrollos, códigos y material generado se encuentra en el repositorio en el portal web <http://ritaportal.udistrital.edu.co:10257/> en donde está alojado el sistema de estimación de riego.

Después de comprender los datos y métodos contenidos en el proyecto se propuso para obtener el objetivo desarrollar la arquitectura propuesta en la figura 4.1.

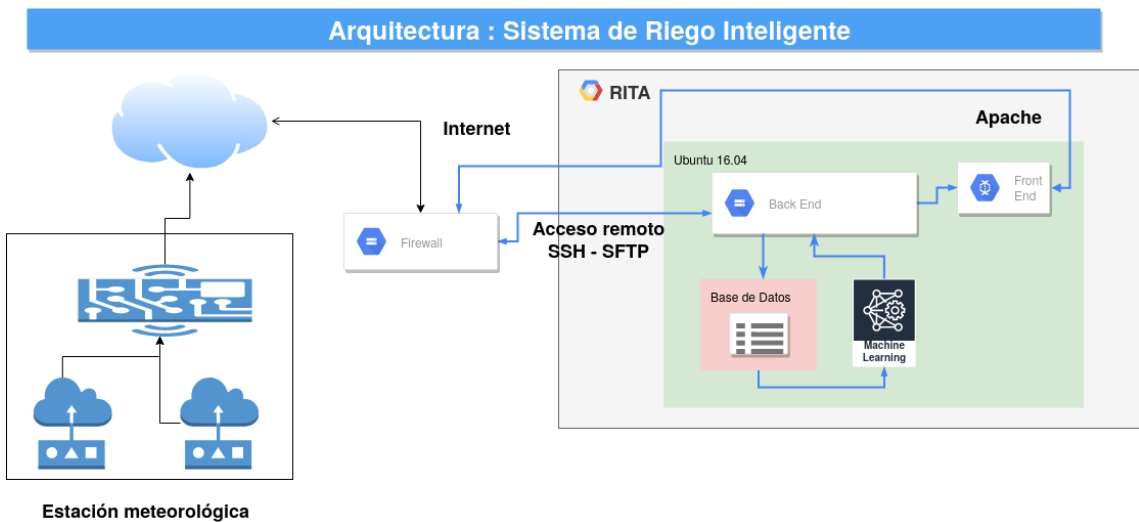


Figura 4.1: Arquitectura propuesta para riego inteligente
Fuente: Propia

Esta propuesta incluye el área de sensado por medio de una estación meteorológica la cual por medio de red GSM transfiere la información a un servidor en la nube el cual presenta la los datos en formato JSON. Estos datos son consultados por la aplicación

desarrollada que los hace pasar por una inteligencia artificial la cual estima el riego necesario a ser aplicado a un cultivo y lo presenta para que el agricultor pueda configurar su sistema de riego.

La implementación del proyecto se realizó siguiendo la metodología de la gestión del conocimiento. Se dio relevancia a los modelos que integran tecnologías destacando el modelo de integración de tecnologías de Kersberg (Díaz, 2005) y el modelo de gestión tecnológica del conocimiento de Paniagua y Lopez (Vanegas, 2015).

Otro tipo de modelos se descartan pues no se centran en gestión de TIC para generación de sabiduría, sino en su uso como herramienta soporte del proceso. Por tener las características de integridad de información y datos necesaria para el proyecto se usa el modelo de Kersberg que propone:

- Asegurar la calidad de los datos e información.
- Transformación de los datos y la información en conocimiento.
- Almacenamiento eficiente y recuperación de metadatos.
- Promoción a tiempo de la diseminación y la distribución del conocimiento.
- Presentación adaptada del conocimiento.

La presente sección se realizó con base a estos postulados siguiendo la estructura presentada en la figura 4.2 para el manejo de la información y los datos pero se estructura respecto a las etapas presentadas en la arquitectura de la solución propuesta en la figura 4.1.

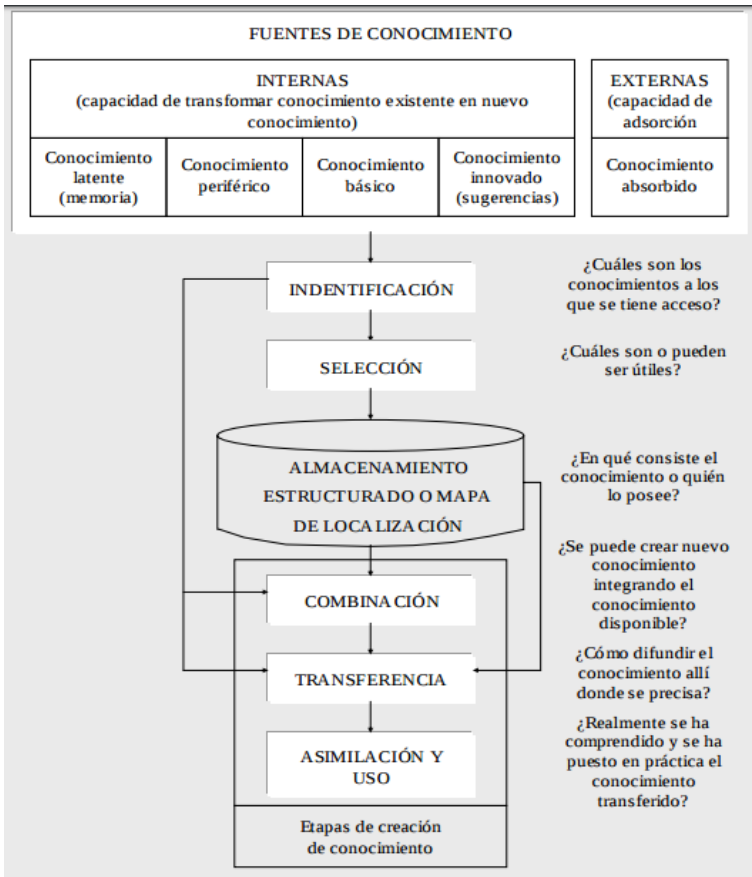


Figura 4.2: Proceso de gestión del conocimiento
 Fuente: (Castillo and Azkuenaga, 2007)

4.1. FUENTES DE CONOCIMIENTO

Esta sección presenta el origen y tratamiento de los datos referenciados en la sección 3, como se integran y cual es su manejo. En la arquitectura desarrollada corresponden a la división *Estación meteorológica* y abarca los datos provenientes de los sensores con sus características y la información externa apropiada por el sistema de riego. De esta forma presenta las variables de interés en las mediciones meteorológicas a ser realizadas, los rangos de estas y los sensores implementados para el levantamiento de datos meteorológicos. El análisis corresponde a la estructura presentada en la figura 4.3.

Las variables se clasifican por el uso que se les da teniendo en cuenta que algunas dan información de primera mano del cultivo o zona a registrar y otras permiten inferir información relevante como la evapotranspiración de referencia en la zona, condiciones apropiadas para crecimiento de material vegetal específico, etc.

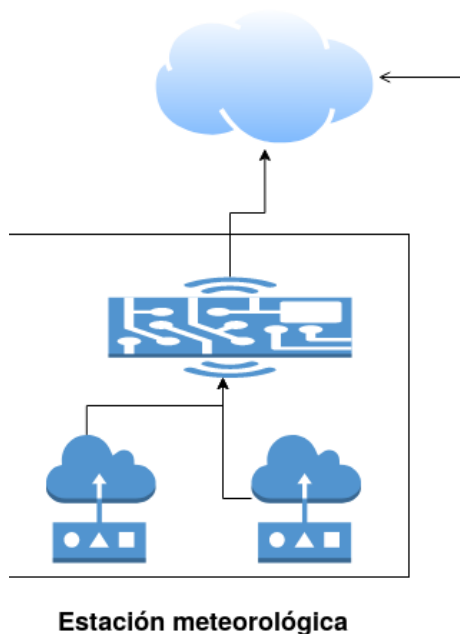


Figura 4.3: Datos meteorológicos y asociados
Fuente: Propia

4.1.1. ESTACIÓN METEOROLÓGICA

TEMPERATURA

La temperatura es una magnitud que mide el nivel térmico o el calor que posee un cuerpo, en este caso las masas de aire que circulan por las zonas a registrar. La medida principal es registrada a dos metros sobre el nivel del suelo y presenta varios valores asociados como:

Temperatura de punto de rocío El punto de rocío es aquel momento en el que condensa el vapor de agua que hay en la atmósfera, así pues la temperatura de punto de rocío es aquella donde se produce este fenómeno.

Índice de calor El índice de calor o sensación térmica es una medida derivada de la temperatura y la humedad que proporciona una descripción de qué forma se percibe la temperatura, este afecta el índice de transpiración de las plantas.

PRECIPITACIÓN

La precipitación relaciona la cantidad de agua que se presenta en una zona por concepto de lluvias, llovizna, nieve, aguanieve o granizo, su medida principal es la tasa de precipitación que corresponde a la cantidad de agua líquida o sólida que alcanza el suelo en cierta unidad de tiempo y su medida es milímetro por hora.

PRESIÓN ATMOSFÉRICA

Esta corresponde al peso de la columna de aire que hay sobre cualquier punto o lugar de la tierra y es por tanto el peso por unidad de superficie, es una medida propia del lugar de referencia y puede ser expresada en hectopascales, pulgadas de mercurio o milibares.

HUMEDAD RELATIVA

La humedad relativa mide la cantidad de agua en el aire en forma de vapor de agua respecto a la cantidad máxima que puede tener la atmósfera a una temperatura dada, se presenta en forma porcentual y debe ser medida a dos metros del nivel del suelo.

RADIACIÓN SOLAR

La radiación solar es una variable importante para el análisis de la evaporación e irrigación, monitoreo de crecimiento de material vegetal, implicaciones en la salud, modelos de predicción del tiempo y clima, entre otros. Corresponde a la energía emitida por el sol y se presenta en vatios por metro cuadrado.

Una variación de la radiación solar es la radiación ultravioleta, es de naturaleza electromagnética comprendida entre los 400 nanómetros y los 15 nanómetros. Forma parte de los rayos solares y permite predecir la forma y color de las plantas, el rendimiento de las cosechas y controlar plagas y enfermedades.

Una medida derivada de la radiación solar es la dosis de radiación, se expresa en MED (Dosis Eritemática Mínima) que es la mínima irradiación necesaria para producir enrojecimiento en la piel tras un determinado tiempo de exposición a la radiación. Se toma como referencia para esta medida sujetos de piel blanca con ligera pigmentación, para contextualizar $1 \text{ MED} = 210 \text{ J/m}^2$.

Se cuenta con otra medida derivada llamada el índice de radiación UV, se expresa en una escala numérica lineal abierta cuyo valor más bajo es 0 e indica el riesgo para una persona de salud normal producido por la exposición a la radiación solar. No hay un valor máximo establecido, aunque más allá de 11 es una situación de riesgo extremo en el que se desaconseja exponerse al sol sin la máxima protección, especialmente entre las 10 y las 16 horas.

VIENTO

La palabra viento refiere el movimiento del aire y se da a consecuencia de la diferencia de presión entre dos puntos. A nivel rural cobra importancia pues transporta polen y materia orgánica, humedad y nubes, masas de aire frío o caliente.

El viento puede tener efectos sobre las superficies por las cuales se desplaza como heladas, causa daños mecánicos a cultivos, erosiona suelos, induce un enraizamiento profundo, deseca el suelo, provoca heladas por evaporación o aumenta la evapotranspiración. Es un fenómeno irregular y siendo una magnitud vectorial los parámetros relevantes son la dirección y velocidad. La dirección indica de dónde viene y la velocidad la relación entre cierta distancia y el tiempo que tarda en recorrerla.

Los datos de interés son la dirección desde la cual proviene en grados, la dirección respecto a los puntos cardinales, la máxima velocidad y la velocidad promedio. Existe una medida de nombre anglosajón referenciada como WIND RUN que corresponde a la cantidad de viento que pasa por un punto en un período de tiempo dado, es decir, la distancia que recorre el viento en una unidad de tiempo.

DATOS GEOGRÁFICOS

Es relevante contar con información geográfica para referenciar las medidas en un punto determinado y conocer las condiciones particulares del terreno, además el tiempo exacto en que son levantados los datos, entre estos se consideran relevantes:

- Estado, provincia o departamento de ubicación de los sensores.
- Municipio o ciudad de ubicación de los sensores.
- País de ubicación de los sensores.
- Coordenadas de latitud y longitud de los sensores.

Para el estimador de riego estos datos son:

Departamento Caldas

País Colombia

Coordenadas Latitud 5.4599237 - Longitud 74.66312408

PROTOCOLO DE COMUNICACIÓN

Cuando la información es agrupada por los sensores se almacenan en formato JSON y se crea un *HOST* que queda activo para escuchar peticiones *HTTP* de clientes para poder presentar esta información a través de GSM a un servidor que presenta los resultados en formato JSON, sin realizar ningún procedimiento de análisis o modificación de los datos. Se puede visualizar este procedimiento en la imagen 4.4.

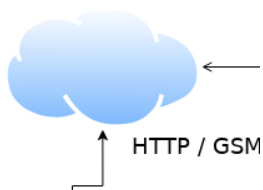


Figura 4.4: Entrada / Salida de información del HOST
Fuente: Propia

4.2. BACKEND - ESTRUCTURA GENERAL

El backend corresponde a la lógica implementada para que la aplicación funcione correctamente. Se trata del conjunto de acciones que ocurren dentro del sistema de riego que no percibe el usuario final pero son fundamentales para su construcción.

En el proyecto el backend se desarrollo en Python y R, para implementar los servicios y el servidor se utilizo el microframework Flask. Corresponde a la sección de la infraestructura y se presenta en la figura 4.5.

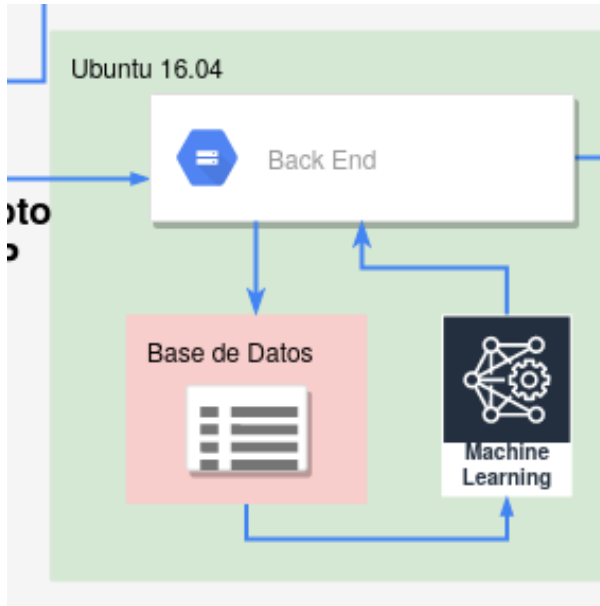


Figura 4.5: Backend implementado
Fuente: Propia

Como se aprecia en la arquitectura, el backend consulta a través de HTTP el servidor donde esta almacenada la información en formato JSON y a través de una consulta GET obtiene la información que procede a guardar en una base de datos SQL, la cual se utiliza para estimar el riego para el cultivo por medio de la etapa de inteligencia artificial, estas secciones se desarrollan a continuación excepto la de inteligencia artificial que tiene un apartado completo de desarrollo mas adelante del documento.

4.2.1. GENERALIDADES

Las variables presentadas han sido etiquetadas y manejadas con criterios de *Web Semántica* para catalogarla en el paradigma de Internet de las Cosas, haciendo que los dispositivos puedan hablar entre ellos sin intervención humana. Para estructurar esta información se revisó el modelo de datos en el lenguaje Web Ontology Language (OWL) en su versión 2.

El lenguaje de ontologías para la web esta diseñado para usarse cuando la información generada o contenida por sensores o documentos necesita ser procesada por apli-

caciones de software. *OWL* se usa para representar explícitamente el significado de términos en vocabulario y la relación entre estos, por eso la representación dada a los datos se adecuan a esta, con el propósito de adecuar el desarrollo a un estándar y permitir futuros trabajos adicionales.

Un nuevo conjunto de especificaciones en 2009 se presentó por la W3C creando *OWL 2* el cual se enfoca en definir ontologías para la web.

4.2.2. ESTRUCTURA DE DATOS

Los datos se estructuran bajo el estándar de la W3C para redes de sensores semánticas, esto con el fin de homogeneizar su tratamiento y tener una estructura común de datos facilitando el proceso del manejo de los mismos. Se hizo uso de la ontología propuesta para redes de sensores agrícolas buscando obtener información de microclimas y datos de plantas. Esta arquitectura permite trabajar con un rango de sensores amplio debido a desarrollos de medición presentes en la industria ya listos para su despliegue, esto hace que el modelo desarrollado se adapte a tecnologías ya existentes y forme la base conceptual para el desarrollo de futuros sistemas de medida.

Este desarrollo se basó en una ontología derivada de la taxonomía de estaciones automáticas de clima presentada en la figura 4.6 y propuesta por la *Organización Mundial Meteorológica* y por el *Climate and Forecast Metadata Conventions NetCDF (Conventions and Metadata, 2018)* quien define los nombres estándar del sistema de clima y pronósticos a nivel mundial a través del programa de diagnóstico e intercomparación del modelo climático que es desarrollado para ser usado con datos climatológicos, de pronóstico, atmosféricos, de superficie y oceánicos.

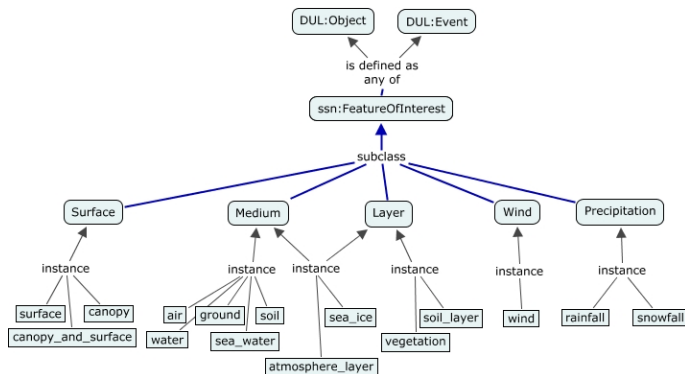


Figura 4.6: Nombres y propiedades relevantes según *clima y pronóstico*

Fuente: (W3C-Incubator-Group, 2011)

En esta figura se presentan las variables de interés para la estimación de parámetros climatológicos relevantes de agricultura definidos como se presentó anteriormente en el lenguaje *DOLCE-Ultralite* parte de *OWL 2*. Puede ser cada medición definida como un objeto o un evento según la necesidad y pertenece en la ontología a características de interés (*ssn:FeaturesofInteres*), las subclases de esta son:

Surface Corresponde a las medidas relacionadas con el suelo directamente.

Medium Corresponde a las medidas relacionadas con las características propias del medio.

Layer Corresponde a las medidas relacionadas con la capa vegetal.

Wind Corresponde a las medidas relacionadas con el viento.

Precipitation Corresponde a las medidas relacionadas con la lluvia.

Así como se presenta la relación de medidas la ontología permite definir el sistema con el grado de detalle que requiramos. El proyecto al ser una arquitectura flexible, agnóstica de tecnología no abarca detalle de cada sensor como sus características físicas pero si como se relacionan las medidas de estos.

La implementación de la ontología, asignando este estándar a las variables presentadas por el sistema de hardware se aprecia en la figura 4.7.

JSON	Raw Data	Headers
Save	Copy	Collapse All Expand All Filter JSON
▼ last_data:		
id:	"098ebfbd-2700-44bb-ab07-0843f7c7e2cd"	
stationid:	"COLCALAZO"	
temp_c:	34.8	
temp_f:	94.6	
relative_humidity:	54	
dewpoint_c:	17	
dewpoint_f:	62.6	
pressure_rel_hpa:	990.8	
pressure_rel_mb:	990.8	
pressure_rel_inhg:	29.3	
pressure_abs_hpa:	990.8	
pressure_abs_mb:	990.8	
pressure_abs_inhg:	29.3	
precip_rate_mmh:	0	
precip_rate_inh:	0	
precip_today_mm:	1.5	
precip_today_in:	0.1	
wind_degrees:	90	
wind_dir:	"E"	
wind_speed_kmh:	5	
wind_speed_mph:	3.1	
wind_speed_ms:	1.4	
wind_gust_kmh:	12.2	
wind_gust_mph:	7.6	
wind_gust_ms:	3.4	
windchill_c:	34.8	
windchill_f:	94.6	
heat_index_c:	41.8	
heat index f:	107.2	

Figura 4.7: Estandarización de nombres de variables según ontología W3C
Fuente: Propia

4.2.3. BASE DE DATOS

La base de datos implementada para el almacenamiento, presentación y estimación de riego tiene la estructura presentada en la figura 4.8 en donde podemos apreciar que

las variables son INTEGER excepto la dirección del viento que es TEXT y la hora del reporte la cual es DATETIME.

Column	Type	Default Value	Nullable	Character Set	Collation	Privileges	Est
id	int(11)	NO	NO			select,insert,update,references	aut
temp_c	int(11)	YES	YES			select,insert,update,references	
relative_humidity	int(11)	YES	YES			select,insert,update,references	
dewpoint_c	int(11)	YES	YES			select,insert,update,references	
precip_today_mm	int(11)	YES	YES			select,insert,update,references	
wind_degrees	int(11)	YES	YES			select,insert,update,references	
wind_dir	varchar(45)	YES	YES	latin1	latin1_sw	select,insert,update,references	
wind_speed_kmh	int(11)	YES	YES			select,insert,update,references	
solar_rad_wm2	int(11)	YES	YES			select,insert,update,references	
report_date_locale	datetime	YES	YES			select,insert,update,references	

Figura 4.8: Base de datos implementada para almacenamiento, presentación y estimación de riego
Fuente: Propia

ALGORITMOS DE CONSULTA Y LIMPIEZA

Parte esencial del Backend es el algoritmo de captura de datos, limpieza de estos y almacenamiento para su posterior análisis, proceso que en la ciencia de datos toma de nombre ETL por las siglas en inglés *Extract, Transformation, Load* cuyo significado es Extracción, Transformación y Carga. El archivo que contiene la implementación de este algoritmo se encuentra con los demás en el repositorio del proyecto. El algoritmo referenciado está desarrollado en el lenguaje de programación R.

El algoritmo como diagrama de bloques se presenta en la figura 4.9 separada por dos colores diferentes, en azul la parte correspondiente a la lectura, captura y limpieza de los datos y en rojo claro la parte correspondiente a conexión con la base de datos y almacenamiento en la misma.

El algoritmo inicia consultando el servidor web que agrega la información capturada por los sensores en formato JSON, del cual extrae los datos de interés e información horaria, este proceso se realiza cada 10 minutos para registrar los cambios en los mismos, la muestreo relaciona cambios brutos en las variables por lo que la frecuencia es acertada.

Luego de este proceso limpia la información estandarizando el formato de fecha y hora para que sea posteriormente entendido por la base de datos, comprueba valores no presentes y por normalización se incluyen. Este proceso concluye al tener la información lista para su almacenamiento.

Para almacenar los datos se abre una conexión con la tabla de datos SQL presentada, en donde previamente se ha configurado el acceso a la misma para lectura, se incorpora la información a una variable que la envía a la base de datos y queda almacenada para su tratamiento posterior.

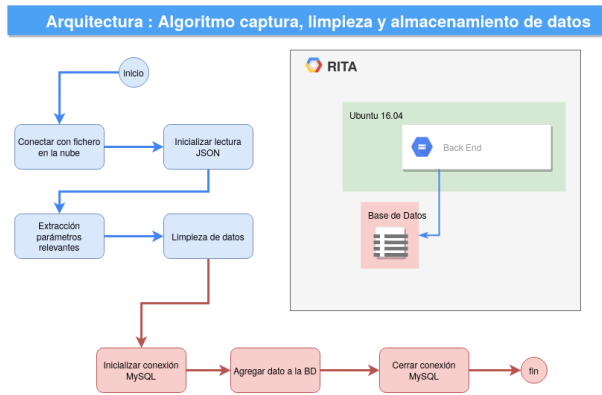


Figura 4.9: Algoritmo de captura, limpieza y almacenamiento de datos
Fuente: Propia

4.2.4. CONFIGURACIÓN DEL SERVIDOR

El servidor corresponde a la máquina que permite cargar los archivos de configuración, la lógica del programa y la visualización del mismo; además de presentar servicios de publicación bajo protocolo HTTP. El servidor es una máquina que contiene Linux Ubuntu como distribución sobre la cual se cargan los paquetes y archivos de configuración para darle salida a Internet.

Corresponde a esta configuración de la salida a Internet el archivo *web.py* el cual inicializa un servidor Flask que maneja la información presente en la base de datos para su presentación.

WEB.PY

Este archivo (*web.py*) monta un servicio de servidor con el microframework de Python "Flask" el cual abre una conexión para presentar el proyecto a través del archivo raíz direccionado a:

<http://ritaportal.udistrital.edu.co:10257/>

Luego de esto establece conexión con la base de datos donde están alojados los datos e inicializa el archivo *index.html* que los renderiza y presenta.

El archivo *web.py* ejecuta una serie de acciones de captura y envío de datos, ejecutando el siguiente procedimiento:

- Se abre conexión con base de datos.
- Se consulta la tabla de sensores.
- Se almacenan en memoria las últimas 50 entradas de información.
- Los datos se organizan para poder ser renderizados.
- Se envían los datos a *intex.html*

De esta forma se articula la base de datos con el Frontend de la aplicación.

CONFIGURACIÓN DE LA MÁQUINA VIRTUAL

La máquina virtual implementada almacena el Frontend, el Backend y los archivos de configuración de la misma así como las rutinas necesarias para su activación.

Para realizar la activación a la máquina virtual se debe entrar por ssh haciendo uso de la llave (la cual se encuentra en los archivos en el repositorio del proyecto) que es un archivo de seguridad *.pem*.

```
ssh -i gabalzatesis.pem servidores@ritaportal.udistrital.edu.co -p 10255 sudo -i R
```

Como librerías de R y Python están en ejecución con diferentes procesos inicia sobre un entorno virtual de python que se encarga de mantener activos los procesos, por eso es necesario cerrar este entorno virtual al ingresar en la terminal por medio de la instrucción.

```
conda deactivate
```

Luego se debe ubicar el usuario *servidores* en la carpeta de usuarios, el cual tiene asignado la carpeta *web*. Al ubicarnos en esta carpeta reconocemos un ambiente virtual el cual activamos con el comando:

```
source env/bin/activate
```

Para luego ejecutar el archivo *web.py*.

```
sudo nohup web.py houp
```

El cual cuenta con la secuencia *nohup*, lo que permite que al cerrar el terminal no se cierre la ejecución del comando.

De esta forma queda activado el servicio web correspondiente sobre el cual se puede ingresar a consultar las variables y el estimado de riego, se profundizará en la presentación de la información en la sección del FrontEnd.

4.3. INTELIGENCIA ARTIFICIAL - CÁLCULO DE LA EVAPOTRANSPIRACIÓN

El cálculo de la evapotranspiración se realiza en dos diferentes fases. Primero se presenta la forma de organizar la información de trabajo para el cálculo directo por medio de la ecuación de Hargreaves que permite la estimación únicamente con parámetros ambientales; en la segunda fase se utilizan estos datos en el tiempo para entrenar modelos de inteligencia computacional y obtener el mejor estimador de riego con los datos anteriormente presentados.

4.3.1. CÁLCULO POR MEDIO DE LA ECUACIÓN DE HARGREAVES

El proceso de cálculo por medio de la ecuación de Hargreaves se desarrolla en dos procesos que son la organización de los datos de trabajo y el proceso mismo del cálculo.

Retomando la ecuación de Hargreaves de la sección 2.2.6, tenemos:

$$ET_0 = 0,0023(T_{mean} + 17,8)(T_{max} - T_{min})^{0,5}R_a$$

Y obteniendo los parámetros ambientales levantados por el sistema meteorológico relevantes como:

- Punto de rocío.
- Índice de calor.
- Precipitación.
- Presión.
- Humedad relativa.
- Fecha de reporte (Dato obtenido por script cada 10 minutos).
- Radiación solar.
- Índice UV.
- Viento en grados.
- Dirección del viento.
- Velocidad del viento.
- Temperatura.

Son estos los datos necesarios para realizar el cálculo de la evapotranspiración de referencia tomando como base la recolección de 13.390 registros en la zona de La Dorada, Caldas; en la franja espacial en latitud 5.4599237 y longitud -74.66312408 y franja temporal entre el 4 de Octubre de 2019 y el 6 de Febrero de 2020.

Para cumplir los requerimientos presentados en la ecuación de Hargreaves se segmentaron estos datos a nivel diario y se calculó de este reporte la siguiente información:

- Temperatura promedio diaria.
- Mayor temperatura diaria.
- Menor temperatura diaria.

Luego de completar la información con estos datos se procedió a ubicar la incidencia de radiación extraterrestre en la zona descrita en los meses de Octubre, Noviembre, Diciembre, Enero y Febrero en la zona según tabla 2.1 para ubicar los valores:

Octubre 15.6

Noviembre 15.6

Diciembre 15.5

Enero 15.6

Febrero 15.8

Complementando los datos para el cálculo de la evapotranspiración de referencia. Así pues con estos datos y usando la ecuación de Hargreaves se computa la necesidad hídrica del terreno restando la pluviosidad presentada al dato calculado.

Esta tabla de entrenamiento se genera en formato csv, se incluye en el repositorio y tiene por nombre *datos_para_entrenamiento.csv*.

Se obtiene un dato presentado en mm/día el cual se recuerda presenta la cantidad de agua a ser usada en un espacio de un metro cuadrado tomando en cuenta que un milímetro de agua equivale a un litro de agua por metro cuadrado que se quiera reponer del cultivo.

Tanto los datos usados para la preparación del experimento como el cálculo de las necesidades hídricas del terreno se encuentran en el repositorio virtual creado para tal fin y puesto a disposición de la comunidad académica para la verificación del mismo bajo conceptos de *investigación reproducible*.

4.3.2. DESARROLLO DEL SISTEMA DE INTELIGENCIA COMPUTACIONAL DE ESTIMACIÓN DE RIEGO

Todos los desarrollos, códigos y material generado se encuentra en el repositorio en el portal web <http://ritaportal.udistrital.edu.co:10257/> en donde está alojado el sistema de estimación de riego.

El cálculo previo de los requerimientos hídricos de un cultivo permite utilizar el aprendizaje supervisado para generar un modelo de inteligencia artificial en el cual las entradas sean el archivo generado para el cálculo de riego (*datos_para_entrenamiento.csv*) presentado en 4.3.1, y los diferentes modelos a entrenar se basen en el apartado 2.4.2 de inteligencia artificial utilizando el 80% de los datos para entrenamiento y 20% para pruebas. El presente reporte presenta el desarrollo de pruebas para determinar el modelo de inteligencia artificial que se aproximara de manera efectiva al comportamiento presentado por los datos meteorológicos de entrada y las necesidades hídricas de salida.

ASOCIACIÓN DE CARACTERÍSTICAS

Como primer se paso se realiza la matriz de asociación que provee información sobre la relación de pares de variables numéricas y categóricas, generando una representación gráfica que por medio de la opacidad del color representa relaciones entre ellas (mas transparente indica una relación más débil y viceversa). Esto sirve para que los diferentes modelos a ser probados predigan agrupamientos o *clusters* que puedan llegar a ser válidos en el desarrollo.

Esta matriz permite:

- Entender la relevancia y naturaleza de la asociación en los datos.
- Detectar asociaciones de pares de variables.
- Identificar clusters de alta asociación de características prioritarias para la construcción del modelo.

La matriz fue creada con el conjunto de datos previamente presentado donde la variable objetivo a predecir es *Riego_mm* que indica la cantidad de agua que requiere el cultivo y se presenta en la figura 4.10.

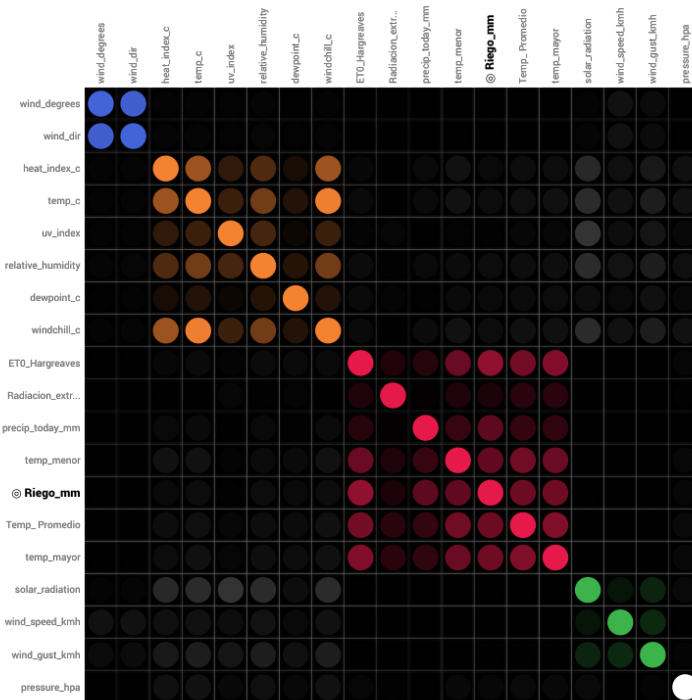


Figura 4.10: Matriz de asociación Fuente: Propia

Para una correcta interpretación debemos saber que:

- Cada punto representa la asociación de dos características, donde la opacidad del color denota la fortaleza de esa asociación.
- Cada cluster es representado por un color diferente.
- La opacidad del color denota el nivel de asociación en un rango de 0 a 1.
- Los puntos grises indican que dadas dos características y que presentan cierto nivel de asociación, no se encuentran en el mismo cluster.
- Los puntos blancos representan características que no se pudieron organizar en ningún cluster.

4

DESARROLLO DEL MODELO

El esquema de trabajo para la comparación de modelos y el desarrollo de comparaciones abarca cuatro pasos que son los siguientes:

- Velocidad de ingesta** Se obtuvo ajustando las variables en el capítulo anterior a su mínima expresión (es el concepto de atomización en las bases de datos relacionales) e identificando la variable objetivo *Riego_mm*, lo cual aumenta la velocidad de tratamiento de los datos.
- Categoría o continua** Si una variable es categórica, la medida de efectividad del modelo obedece a diferenciar la categoría y el objetivo es encontrar un modelo que minimice la función estadística *LogLoss*, si la variable es continua como es la predicción de riego se pueden usar tres criterios estadísticos de regresión tales como *RSME*, Error de la Raíz Media Cuadrática (Por sus siglas en inglés); *R-Squared* o R-cuadrada que depende directamente de la varianza de la variable dependiente y el test general F el cual es una prueba estadística bajo una hipótesis nula que permite comparar modelos estadísticos. Todos basados en dispersión de raíces. El proyecto utiliza *RSME* por ser el que más información acumula.
- Segmentación** Los modelos toman los conjuntos de datos preparados y preprocesan la información además de segmentarla para su uso en ocasiones por la técnica de *map-reduce*.
- Entrenamiento** Con estos parámetros preestablecidos se entrenan los diferentes modelos para verificar el de mejores características de respuesta.

PRESUNCIONES DEL MODELO

Los métodos de inteligencia computacional pueden producir modelos predictivos más precisos que los tradicionales como la regresión estadística, pues estos últimos se basan en probar una hipótesis base para la ponderación de cada variable y la selección de características probando diferentes tipos de errores (*ejemplo t-test, p-value, error estándar*) donde muchas de estas presunciones pueden no ser soportadas por los datos. Los métodos de inteligencia computacional ofrecen mayor flexibilidad en la estructura del modelo, los cuales típicamente tienen un mayor rendimiento pues no se basan en una hipótesis y tienen una selección de características internas de manera heurística.

Una limitante común de los modelos de aprendizaje de máquina es el sobre entrenamiento, el cual ocurre cuando no se realiza una estimación o predicción, sino que el sistema solo imita la salida deseada, derivando en un modelo incapaz de dar una respuesta satisfactoria ante una entrada diferente a las que fue entrenado. Para evitar llegar a este punto de sobreentrenamiento se realiza una verificación en cada momento del entrenamiento llamada validación cruzada, *Cross-Validation* para asegurar la validez del modelo. Lo que hace es segmentar tanto los datos de entrenamiento como los de prueba en diferentes secciones para sucesivamente ir calculando la media aritmética y comprobando que sea homogénea a lo largo de toda la distribución.

SELECCIÓN DE VARIABLES

La lista de variables a usar debe de preferencia tener información que contribuya al ajuste y despliegue del modelo, teniendo en cuenta que mas listados de variables crearan un modelo de mayores proporciones que actuará con una velocidad menor; caso contrario no se incluyan variables relevantes se perderá exactitud del mismo.

Entre las variables que no contengan información se relacionan aquellas que estén conformadas por ejemplo por solo unos, o solo ceros, columnas duplicadas o una que sea función lineal de otra (en ese caso no estaría añadiendo información al modelo).

Para encontrar esa relación verificamos la carga estadística de cada variable con su nombre, número de valores únicos, número de valores inexistentes, media, desviación estándar, mediana, valor mínimo y máximo. La información se presenta en la tabla 4.1.

Nombre	Único	Inexistente	Media	DesvEst	Mediana	Mín	Max
precip_today_mm	96	0	1.86	5.57	0.0	0.0	66.3
wind_degrees	16	0	151.17	71.55	135.0	0.0	338.0
wind_dir	16	0	N/A	N/A	N/A	N/A	N/A
windchill_c	194	0	29.63	4.0208	28.8	21.6	41.5
temp_c	194	0	29.64	4.022	28.8	21.6	41.5
ET0_Hargreaves	117	0	5.7	0.57	5.78	3.78	7.3
Riego_mm	413	0	3.89	5.67	5.64	-60.089	7.3

Cuadro 4.1: Carga estadística de las variables. Fuente propia

Se aprecia que la variable *wind_degrees* no responde a los valores de media, desviación estándar, mediana, mínimo o máximo pues es una variable categórica que relaciona la dirección de desplazamiento del viento, siendo norte, sur, este, oeste y sus variaciones intermedias noreste, sureste, etc.

RELACIONES DE DATOS

Para obtener la relevancia que tiene cada variable se cambian datos o hasta tipos de datos comprobando la relación lineal o estadística entre variables. Técnica llamada *Importancia de Permutación* que mide cuán tanto empeora la predicción o estimación modificando una columna para luego normalizar ese valor y presentar las relaciones que tiene cada variable con la posible generación de un modelo.

En el desarrollo del modelo se realizó la prueba "Importancia de Permutación" donde el resultado obtenido se presenta en la tabla 4.2 y en la figura 4.11 donde se muestra luego de ser normalizado.

Feature Name	Impact Normalized	Impact Unnormalized
precip_today_mm	1.0	7.1815
ET0_Hargreaves	0.1145	0.8222
wind_degrees	0.1113	0.7991
wind_dir	0.1113	0.7991
windchill_c	0.0504	0.3622
temp_c	0.0504	0.3621

Cuadro 4.2: Resultados de la prueba "Importancia de permutación" totales y normalizados. Fuente propia

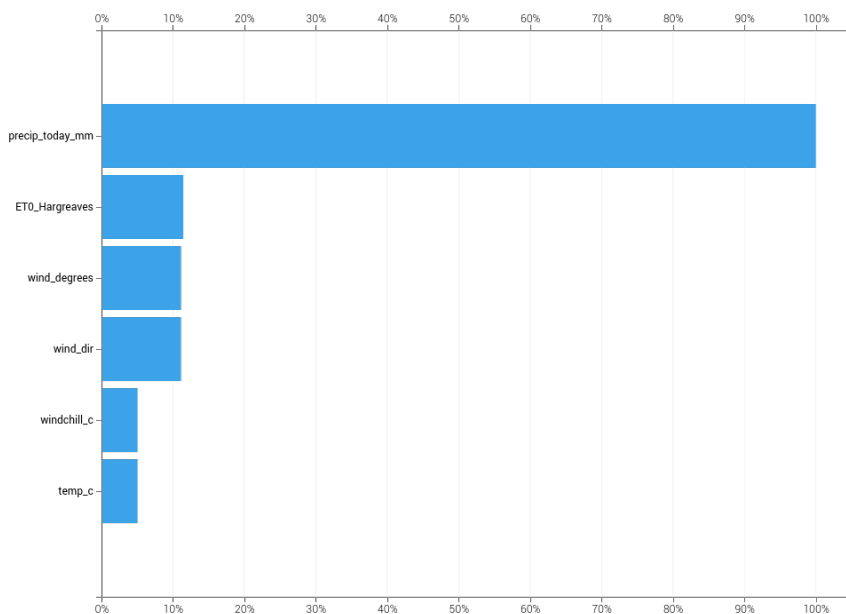


Figura 4.11: Importancia de Permutación normalizada Fuente: Propia

Como se aprecia la variable que tiene directa relación con los valores que pueda llegar a estimar el modelo desarrollado es la precipitación, esto debido a que es un factor directo en la cantidad de riego que necesita el cultivo, y si llueve la medida necesaria, no se debe agregar mas agua al cultivo.

ENTRENAMIENTO DE MODELOS

En orden de encontrar el modelo adecuado para la estimación de riego se proceden a validar aquellos centrados en manejo variables continuas basados en redes neuronales, árboles de decisión, gradientes de descenso, *RandomForest*, regresores y máquinas de soporte vectorial.

Para realizar este proceso se debe asegurar en primer instancia no sobre-entrenar el modelo realizando una validación cruzada en aquellos que muestren un resultado satisfactorio. El proceso se realizó de la siguiente manera:

- Se separa el 20 % de los datos como conjunto de prueba.
- Los datos restantes son separados en 5 conjuntos de validación y se inicia el entrenamiento de cada modelo con estos diferentes subconjuntos.
- Aquellos modelos con un alto desempeño se reentrenan con los conjuntos adicionales mientras sigan presentando características de alto desempeño.
- De esta forma solamente los modelos con mayor rendimiento se entrenan con la totalidad de los datos, esto con el fin adicional de aumentar el rendimiento de la máquina pues la validación de 30 conjuntos es un trabajo que exige altas prestaciones.
- Los modelos con mayor rendimiento son comparados según la validación cruzada y se concluye cual es el más acertado.
- Se comienza a validar el desempeño del modelo y a través de la calidad de la estimación que realiza decidir si cumple los requisitos o se cambia de modelo.

Cabe recordar que la métrica usada para la validación inicial como para la validación cruzada es el error de la raíz cuadrática media (*RMSE*).

RESULTADOS DEL ENTRENAMIENTO

Los diferentes modelos entrenados con sus respectivos valores de validación se presentan en la tabla 4.3, 4.4 y 4.5. Dando como resultado que el modelo a utilizar para predecir el riego en un cultivo con los datos de entrenamiento y parámetros de entrada es Light Gradient Boosting on Elastic-Net Predictions.

El modelo Boosting se basa en pequeños estimadores que uniéndolos realizan una estimación de mayores rendimientos y que sea de gradiente ligero (*light gradient*) implica un modelo rápido, distribuido y de alto rendimiento basado en árboles de decisión.

Dado que el modelo se basa en algoritmos de árbol de decisión, divide la rama del árbol con el mejor ajuste (por *RMSE*), mientras que otros algoritmos de refuerzo dividen la profundidad del árbol o el nivel en lugar de las ramas. Así es como este modelo puede reducir mayores pérdidas que el algoritmo de nivel y, por lo tanto, da como resultado una precisión mucho mejor. La velocidad con que realiza este proceso le da su nombre de ligero.

Modelo	Validación	V. Cruzada	Muestra
Light Gradient Boosting on Elastic-Net Predictions	0.002	0.0021	100
GLM Blender	0.001	0.0009	63.9928
Advanced GLM Blender	0.001	0.0009	63.9928
Advanced AVG Blender	0.0047	0.004	63.9928
AVG Blender	0.0054	0.0055	63.9928
Elastic-Net Regressor (mixing alpha=0.5 / Least-Squares Loss)	0.0104	0.0107	63.9928
ENET Blender	0.0118	0.0108	63.9928
Linear Regression	0.0184	0.017	63.9928
Elastic-Net Regressor (mixing alpha=0.5 / Least-Squares Loss) with Unsupervised Learning Features	0.0205	0.0213	63.9928
Vowpal Wabbit Regressor	0.0249	0.013	63.9928
Nystroem Kernel SVM Regressor	0.0347	0.0857	63.9928

Cuadro 4.3: Diferentes modelos entrenados y considerados

Modelo	Validación	V. Cruzada	Muestra
Ridge Regressor	0.0478	0.05	63.9928
Gradient Boosted Trees Regressor (Least-Squares Loss)	0.0603	N/A	31.9964
Decision Tree Regressor	0.0739	N/A	31.9964
Gradient Boosted Greedy Trees Regressor with Early Stopping (Least-Squares Loss)	0.0939	N/A	31.9964
RuleFit Regressor	0.1298	N/A	31.9964
RandomForest Regressor	0.1628	N/A	31.9964
Gradient Boosted Trees Regressor	0.175	N/A	31.9964
Eureqa Generalized Additive Model (1075 Generations)	0.1848	N/A	31.9964
Gradient Boosted Trees Regressor with Early Stopping(Least-Squares Loss)	0.271	N/A	15.9982
Breiman and Cutler Random Forest Regressor	0.2874	N/A	15.9982

Cuadro 4.4: Diferentes modelos entrenados y considerados parte 2

4.3.3. ANÁLISIS DEL MODELO DESARROLLADO LGB ON ELASTIC-NET

Cada modelo de inteligencia artificial tiene un desarrollo por diagramas de bloques el cual puede explicar de una manera gráfica el flujo de trabajo del mismo desde el momento en que se ingestan los datos, la separación en variables numéricas y categóricas,

Modelo	Validación	V. Cruzada	Muestra
eXtreme Gradient Boosted Trees Regressor with Early Stopping	0.2918	N/A	15.9982
Light Gradient Boosted Trees Regressor with Early Stopping	0.2945	N/A	15.9982
eXtreme Gradient Boosted Trees Regressor with Early Stopping and Un-supervised Learning Features	0.2985	N/A	15.9982
ExtraTrees Regressor	0.302	N/A	15.9982
Generalized Additive2 Model	0.3388	N/A	15.9982
TensorFlow Neural Network Regressor	0.5775	N/A	15.9982
Auto-tuned K-Nearest Neighbors Regressor (Euclidean Distance)	1.1314	N/A	15.9982
Ridge Regressor with Binned numeric features	2.8336	N/A	15.9982
Mean Response Regressor	5.1192	N/A	15.9982
Isolation Forest Anomaly Detection	6.449	N/A	15.9982

Cuadro 4.5: Diferentes modelos entrenados y considerados parte 3

la limpieza de los datos y el trabajo que se describió anteriormente.

La descripción del modelo implementado por medio de diagramas de bloques, se visualiza en la figura 4.12, la cual presenta el flujo de trabajo de la información desde el momento que es apropiada hasta cuando realiza la estimación.

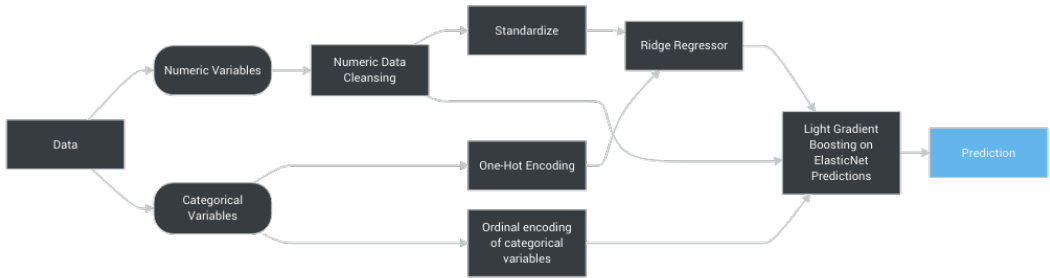


Figura 4.12: Flujo de Trabajo del modelo Light Gradient Boosting Fuente: Librería TensorFlow

Se describe a continuación el proceso de:

- Limpieza de datos numéricos
- Regresor de árboles de decisión Boosted con gradiente ligero y parada temprana
- Codificación a un estado

- Estandarización
- Transformación de variables categóricas a ordinarias
- Regularización por Elastic Net

LIMPIEZA DE DATOS NUMÉRICOS

El modelo al digerir datos realiza un proceso de verificación que garantice el no ingreso de estructuras de datos con valores faltantes o datos de baja calidad, esto refiere a que los datos de entrada no cambien de formato, combinen puntos o comas de forma inesperada, la aplicación o un exceso de ceros donde no es común u otro tipo de errores de ingreso de datos que puedan llegar a crear un mal funcionamiento en la predicción.

4

REGRESOR DE ÁRBOLES DE DECISIÓN BOOSTED CON GRADIENTE LIGERO Y PARADA TEMPRANA

El Regresor de Gradiente Ligero con Parada Temprana usa la implementación de LightGBM de gradientes de arboles de decisión. Este utiliza las perdidas de mínimo cuadrado de forma predeterminada y funciona de la siguiente manera.

Las máquinas de aumento de gradiente o *Gradient Boosting* son técnicas de inteligencia artificial para problemas de regresión y clasificación que producen un modelo de predicción ensamblando pequeños modelos de decisión, típicamente *árboles de decisión*. Los algoritmos que contiene este modelo realizan un entrenamiento extremo y son considerados de los modelos mas versátiles que existen pues da un balance eficiente entre tendencia (*bias*) y la varianza.

Son muy similares al método *Random Forest* que entrena árboles de decisión individuales para luego reentrenar conjuntos de ellos donde cada árbol continua siendo un punto particular del modelo. *GBM* a diferencia de *Random Forest* entrena los árboles de decisión en paralelo y luego entrena unos árboles con el error combinado de los anteriormente entrenados para así crear un lazo de corrección de error. Este modelo debe balancear de forma preferencial la tasa de aprendizaje y el número de arboles en el modelo.

La sección del modelo de parada temprana (*Early Stopping*) se incluye para determinar el menor número de árboles posible para realizar una estimación acertada antes que el modelo comience a sobre entrenarse.

CODIFICACIÓN A UN ESTADO

Este proceso parte de las propiedades de transacciones en bases de datos que busca garantizar la integridad de las mismas (Principio ACID - *atomicity, consistency, isolation, durability*) buscando la atomicidad del dato.

Este paso en el desarrollo del modelo crea una codificación binaria. Un valor booleano es construido para cada posible cadena que pueda existir dentro de los arreglos de datos, de esta forma los cálculos se realizan de una manera mas rápida y las tablas de datos se llevan a su mínima expresión, este procedimiento es relevante con entradas categóricas. Para aquellas con dos únicos valores solo se crea un valor booleano.

ESTANDARIZACIÓN

Este paso en el modelo refiere a la normalización de las variables para que los datos que contienen sean acordes en magnitud entre las diferentes variables y alguna que presente datos de magnitudes diferentes (cientos contra miles) no agreguen esta información de magnitud al modelo, haciendo que todas tengan el mismo rango. Se estandarizan los datos removiendo la mediana y escalando los resultados con la desviación estándar.

Este paso es comúnmente requerido en múltiples modelos de inteligencia de máquina pues muchos elementos usados en el algoritmo de aprendizaje asumen que las variables están centradas en cero y que su varianza es del mismo orden; si alguna variable tiene una varianza en mayores ordenes de magnitud que las demás esta dominara la función objetivo y hará que el estimador no aprenda las características como se espera.

TRANSFORMACIÓN DE VARIABLES CATEGÓRICAS A ORDINARIAS

Este paso convierte variables categóricas en ordinarias dependiendo la cantidad de estados que pueda poseer, en este caso particular se presenta la variable de dirección del viento la cual contiene información cardinal.

REGULARIZACIÓN POR ELASTIC NET

ElasticNet es un modelo de regresión regularizada que combina penalidades L1 y L2 que se usan cuando los conjuntos de datos tienen gran cantidad de variables. L1 es llamada *Lasso Regression* y L2 *Ridge Regression* en donde L1 agrega el valor absoluto de la magnitud de error como penalidad a la función de pérdidas y L2 el cuadrado de esta magnitud.

El modelo realiza este procedimiento cuando se encuentran variables de alta correlación como la pluviosidad con el riego.

4.3.4. EVALUACIÓN DE LA ESTABILIDAD DEL MODELO

Para corroborar la pertinencia del modelo se realiza una *Tabla de Elevación* o *Lift Chart*. Esta organiza y agrupa características numéricas en segmentos de igual tamaño para realizar la estimación con los datos presentes y evaluar que tan precisa es la predicción, lo que ayuda a evaluar la precisión del modelo en cada sección de datos diferentes para ver cuán susceptible a ser mejorada o corregida.

Esta se organiza por valores predichos de menor a mayor. La gráfica con la tabla de elevación del modelo se presenta en la figura 4.13.

Los puntos en la tabla de elevación representan el porcentaje promedio de predicción en cada sección de datos analizada. Los símbolos azules de suma indican el valor promedio de la predicción para el conjunto de datos y los círculos naranja el valor promedio de los datos en esa sección. Como se puede apreciar existe una cercana relación entre las dos curvas lo que infiere un comportamiento eficiente del modelo con un alto nivel de exactitud.

El modelo desarrollado toma los datos contenidos en las variables de entrada y realiza una estimación del riego necesario para un terreno de un metro cuadrado de pasto, el cual es por definición la métrica calculada al inferir la evapotranspiración de referencia (*ET_o*).

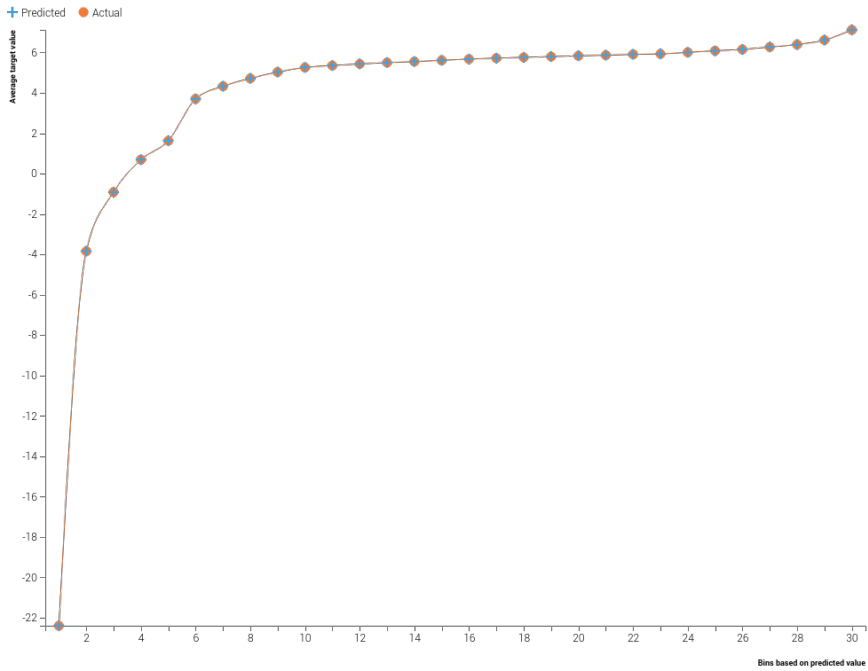


Figura 4.13: Tabla de elevación del modelo desarrollado Fuente: Keras

Para identificar las necesidades hídricas de un cultivo en particular se procede con este estimado a asignar una medición del terreno a ser cubierto por el riego y el tipo de cultivo. Estas dos variables influyen de manera directa en el proceso pero son susceptibles a cambiar en el tiempo por la asignación del terreno a otro cultivo o por el cambio en el tamaño del sector a regar.

4.4. FRONTEND - DASHBOARD

Todos los desarrollos, códigos y material generado se encuentra en el repositorio en el portal web <http://ritaportal.udistrital.edu.co:10257/> en donde está alojado el sistema de estimación de riego.

El dashboard se desarrolló integrando un CSS y código javascript dentro de la plantilla de código libre *Alpha by HTML 5 UP* la cual permite modificar su landing page para incorporar puntos de acceso a información. El proyecto se encuentra dentro de la carpeta *web/front* y el código principal esta consignado en el archivo *index.html* que recibe la información organizada por el Backend y la presenta en el explorador web.

4.4.1. INDEX.HTML

Este archivo maneja la presentación de la información de la siguiente manera:

- Configura la codificación en utf-8 para que el sistema reconozca caracteres latinos.
- Configura una estructura donde se presenta un banner, un contenedor de información meteorológica y un contenedor de datos del proyecto.
- El contenedor de información meteorológica presenta los datos de las variables ambientales en tiempo real.
- Se presenta junto a estos datos la medida del riego a ser suministrado al cultivo.

4.4.2. IMÁGENES DEL SISTEMA



Figura 4.14: Banner del sistema
Fuente: Propia

Un manejo adecuado de los recursos de agua representará un aumento más importante en los rendimientos de los cultivos agrícolas que el conjunto de todas las otras prácticas de producción juntas.

FAO - Naciones Unidas

Riego en milímetros



44 mm

Figura 4.15: Encabezado del sistema
Fuente: Propia

Un manejo adecuado de los recursos de agua representará un aumento más importante en los rendimientos de los cultivos agrícolas que el conjunto de todas las otras prácticas de producción juntas.

FAO - Naciones Unidas

Temperatura



37 °C

Esta es una medida principal, de ella deriva la temperatura media, la máxima y la mínima diaria para entrenar el modelo de inteligencia artificial.

Humedad



44 %

La humedad indica el grado de saturación de vapor de agua en las cercanías del cultivo.

Figura 4.16: Variables del sistema
Fuente: Propia

4.5. DOCUMENTO TÉCNICO DE USO DEL SISTEMA DE RIEGO

El presente apartado presenta un documento técnico que usa la gestión del conocimiento para integrar las diferentes variables; su interrelación permite mejorar la forma en que se controla el riego en las granjas inteligentes en Colombia.

Como se presentó en la sección "Desarrollo", se organizó el proceso de gestión del conocimiento desde la propuesta de Castillo para asimilar y utilizar conocimiento almacenado y generado por fuentes digitales (Castillo and Azkuenaga, 2007). El proceso se presenta en la figura 4.19.

Partiendo de este concepto procederemos a identificar:



Figura 4.17: Variables del sistema
Fuente: Propia



Figura 4.18: Pie de página del sistema
Fuente: Propia

- Fuentes de conocimiento.
- Identificación de variables.
- Selección de variables relevantes.
- Almacenamiento estructurado.
- Combinación y transferencia.
- Asimilación y uso.

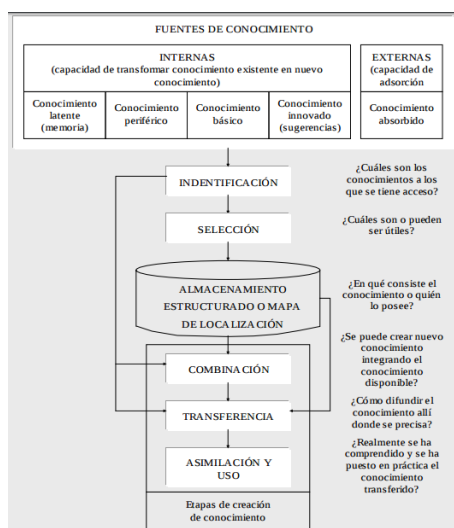


Figura 4.19: Proceso de gestión del conocimiento

Fuente: (Castillo and Azkuenaga, 2007)

4.5.1. FUENTES DE CONOCIMIENTO

En primera instancia identifique sus fuentes de conocimiento referenciando si son internas o externas. Para el propósito de uso del sistema de riego las variables de la estación meteorológica son de uso interno pues ya están contenidas en la aplicación, si se desea vincular una nueva estación es necesario conocer el protocolo de comunicación de la información.

Si la comunicación esta codificada en formato JSON o CSV la implementación en el código es transparente y equivaldría a duplicar las entradas en la base de datos para llevar un doble registro temporal y espacial. Para llevar a cabo este procedimiento comuníquese con el administrador del sistema para que la misma sea integrada y se le genere un dashboard particular.

Las variables climáticas de fuentes externas como páginas web o sistemas agroclimáticos independientes ya están contendidas dentro del sistema por deducción de las principales.

Si considera alguna variable adicional que crea le es de interés comuníquese con el administrador del sistema para validar su incorporación.

4.5.2. IDENTIFICACIÓN DE VARIABLES

Las variables que se han tenido en cuenta para la estimación de riego son las contenidas dentro del proceso de transferencia de información, estas incluyen:

- Punto de rocío.
- Índice de calor.
- Precipitación.

- Presión.
- Humedad relativa.
- Fecha de reporte (Dato obtenido por script cada 10 minutos).
- Radiación solar.
- Índice UV.
- Viento en grados.
- Dirección del viento.
- Velocidad del viento.
- Temperatura.

4.5.3. SELECCIÓN DE VARIABLES RELEVANTES

Las variables descritas anteriormente son de interés, pero por el grado de información que contienen se priorizan algunas de ellas, modificando la estimación de manera mínima. Las variables prioritarias que contienen la mayor cantidad de información se describen en el apartado de selección de variables, y pueden ser consultadas en la tabla 4.1 con la carga estadística que contienen; se relacionan a continuación:

- Precipitación.
- Humedad relativa.
- Fecha de reporte (Dato obtenido por script cada 10 minutos).
- Radiación solar.
- Viento en grados.
- Dirección del viento.
- Temperatura.

4.5.4. ALMACENAMIENTO ESTRUCTURADO

El almacenamiento de la información se realiza sobre una base de datos estructurada SQL que contiene los patrones de lectura habilitados para que pueda ser leída por múltiples sistemas de consulta de manera simultánea, mientras se disponga de las claves de la base de datos.

La tabla donde se almacena la información tiene por nombre *consultasensores* y comparte este mismo nombre con la contraseña *consultasensores*. Si se quiere incluir más información a la tabla es necesario agregar las mismas en lenguaje SQL.

4.5.5. COMBINACIÓN Y TRANSFERENCIA

La información al salir de la base de datos ingresa al sistema de inteligencia artificial sobre el cual el usuario no puede realizar modificaciones, al terminar se transfiere al frontend para ser presentada.

4.5.6. ASIMILACIÓN Y USO

Ya luego de tener esta información corresponde al agricultor determinar según el tamaño en metros cuadrados del cultivo y tipo de cultivo cual será la cantidad de agua a agregar con base en los siguientes parámetros:

- El sistema entrega la estimación riego en mm, esta medida corresponde a la cantidad de agua que necesita un cultivo de pasto de un metro cuadrado de diámetro.
- Cada mm de agua significa un litro a ser usado en ese terreno descrito.
- Así pues, si el terreno tiene por dimensiones 20 metros de frente por 3 de fondo significa que tiene 60 metros cuadrados y este 60 será el valor por el que hay que multiplicar el riego en mm para saber cual es la necesidad hídrica del terreno.
- El valor hallado previamente corresponde al pasto, cada cultivo tiene una necesidad hídrica diferente que se realiza por medio del cálculo $ET_c = ET_o * K_c$ como se presento en el apartado *Factores que afectan la ET* y en la figura 2.3.
- Al ser K_c un factor lineal se puede realizar la multiplicación no con ET_o sino con el valor de *Riego en mm* y escalarlo de forma automática a las dimensiones necesarias del cultivo referenciando los valores de la tabla 4.6.
- Esta tabla cuenta con tres opciones, la primera indica el valor por el cual debe multiplicarse el riego en el periodo de crecimiento inicial de la planta, el segundo el valor de crecimiento a la mitad de su tiempo de cosecha, y el último al finalizar la etapa de crecimiento.

Grupo de cultivo	Kc ini	Kc med	Kc fin
Hortalizas pequeñas	0.7	1.05	0.95
Tomate, chile, pimiento	0.6	1.15	0.80
Pepino, melón, sandía	0.5	1	0.80
Raíces y tubérculos	0.5	1.10	0.95
Frijol, habichuela, arveja	0.4	1.15	0.55
Alcachofa, fresas	0.5	1	0.80
Ajonjolí, girasol	0.35	1.15	0.35
Cereales	0.3	1.15	0.40
Caña de azúcar	0.4	1.25	0.75
Banano	0.5	1.10	1
Piña	0.5	0.30	0.3

Cuadro 4.6: Kc para diferentes tipos de cultivo en diferentes etapas de crecimiento. Fuente: FAO (FAO, 2017)

Así pues, es como se debe operar el sistema de estimación de riego.

5

CONCLUSIONES

La planta devuelve al medio el 90% del agua que recibe, por lo que controlar la cantidad de agua que tenga disponible es importante para su correcto crecimiento, el avance de la tecnología brinda herramientas importantes para organizar y tomar control de procesos los cuales de forma clásica han sido abordados por soluciones de alto costo, lo que permite desarrollar sistemas cada vez mas asequibles a diferentes tipos de productores y con mayor eficiencia.

Para determinar la evapotranspiración (*ET*) experimentalmente se requieren instrumentos específicos y mediciones precisas de varios parámetros físicos. Generalmente los métodos aplicados en zona rural son caros y solo realizados apropiadamente por personal de investigación con un alto nivel de preparación. El método propuesto ha resultado ser eficiente y necesita tan solo un sistema de hardware que levante información ambiental en el sitio que se quiere calcular el riego, esto puede ser suplido por un diseño propio, independiente o con una estación meteorológica.

La planta necesita un sistema de riego que garantice la disponibilidad de agua cuando no hay presencia de lluvias en el cultivo para obtener un crecimiento efectivo, sin caer en deficiencia ni en saturación de agua y para realizarlo de manera eficaz se necesita determinar factores edafológicos (suelo - clima) además la disponibilidad de agua en la zona a implementar el cultivo.

Se utiliza para la medición de ET el método propuesto por la FAO ya que un sistema de gestión de conocimiento es adaptable a ser alimentado por variables meteorológicas provenientes de sensores ubicados en puntos estratégicos del cultivo, lo que va en línea directa con la propuesta desarrollada.

El objeto principal para desarrollar un proceso de inteligencia artificial es el *dato*, que representa un objeto o evento observable; cuando se reúne y ordena una serie de datos

se eleva el nivel de conocimiento generando *información* en la que ya es posible encontrar significado y llegar al *conocimiento*. Este fue el procedimiento seguido y presentó un desarrollo fiable de la pregunta de investigación.

El proyecto tiene limitantes geográficas debido a su instalación en zonas aisladas del territorio rural donde no se presentan sistemas de conectividad adecuados. La propuesta para solucionar este percance es utilizar una alternativa de código abierto que pueda ser instalada de forma local sin necesidad de salir a la nube, creando un ecosistema de dispositivos locales según el paradigma de la computación de techo.

En cuanto al despliegue de tecnologías de la información y las comunicaciones en el sector agrícola Colombia se halla en un punto anterior al de las potencias agrícolas mundiales, no por incapacidad del productor nacional sino por las condiciones económicas imperantes y las condiciones orográficas nacionales.

Las industrias de tecnología agrícola desarrollan sus productos para suplir la demanda de escala que refiere a terrenos planos de altura estándar y condiciones homogéneas a lo largo del periodo de cultivo, cosas que no se presentan en Colombia pues los cultivos que necesitan ser automatizados quedan en laderas, los afectan diferentes tipos de climas y no son homogéneos en el tipo de suelo que contienen, lo que hace difícil para una empresa desarrollar toda una línea tecnológica solo para el mercado nacional.

El desarrollo realizado es apto de implementarse a nivel mundial, es la base de una tecnología de alto impacto. Sin embargo para afinar el modelo de inteligencia artificial sería conveniente alimentarlo con información de cantidades de producción a nivel local para poder estimar no solo necesidades hídricas sino proyecciones de producción, lo cual sería muy valioso para modelos económicos como la compra de bonos verdes de contaminación y la mejora de las cadenas de valor agrícolas.

6

TRABAJOS FUTUROS

6.0.1. TRABAJOS EN INGENIERÍA ELECTRÓNICA

El desarrollo de hardware relacionado con procesos de medición meteorológica ha tenido grandes avances en los últimos años, los precios han bajado y acercan estas tecnologías a las masas. Sin embargo es necesario crear procesos de hardware modular que pueda ser implementado por secciones según las condiciones del agricultor para dar procesos de crecimiento tecnológico continuo a un cultivo, este tipo de sistemas es necesario a nivel nacional y no existen alternativas que lo resuelvan, por lo que lo convierte en un proyecto de grado con bastante utilidad.

De forma reciente se encuentran en el mercado plataformas de hardware que incorporan procesos de inteligencia artificial, optimizados para ciertos modelos; un caso de estos es el Intel Corporation Neural Compute Stick 2s. Es de interés conocer el manejo de esta arquitectura para desarrollar procesos integrados de hardware + AI como el desarrollado en esta tesis, podrías ser un paradigma el cual cambie el proceso de manejo de información y acceso a periféricos de aplicaciones.

Es de interés realizar un trabajo similar al propuesto en esta tesis pero que no tenga conexión a Internet, realizando el cálculo del riego de forma local. Esto pues en zonas rurales apartadas de núcleos urbanos la conectividad aún de una red celular es nula. Así pues un sistema que realice esto de forma local bajo el paradigma de *Computación de techo* es un desarrollo que vale la pena explorar.

6.0.2. TRABAJOS EN INGENIERÍA DE SISTEMAS

El desarrollo del proyecto implica el manejo de bajos volúmenes de datos, lo que no permite escalar la aplicación hacia el paradigma de Big Data, ni usar estrategias de Map Reduce para analizar estos datos en gran cantidad, sin embargo estas variables meteorológicas aumentaran su cantidad de registros paulatinamente y aumentara la cantidad de las mismas así que se generara información con una velocidad exponencial.

En este punto se hace necesario el desarrollo o adecuación de una plataforma que pueda soportar grandes volúmenes de datos y manejar inteligencia artificial sobre ellos, lo que sería un buen proyecto de pregrado.

Con los datos disponibles se pueden próximamente entrenar modelos cada vez más complejos de inteligencia artificial; ya teniendo los datos y el procedimiento descrito en el documento se puede realizar de manera mecánica, e ir probando variaciones nuevas que se vayan creando de los modelos.

La ontología usada para la realización del proyecto está desarrollada en XML, sería un proyecto interesante transferirla a JSON para los dispositivos hardware que ya usan este formato, pues permitiría integrarlos al estándar y habilitarlos para transferir información de manera transparente hacia cualquier sistema de analítica del mundo.

6.0.3. TRABAJOS EN INGENIERÍA INDUSTRIAL

En el desarrollo del proyecto se encontraron otros procesos dentro de la cadena agrícola propensos a ser estudiados y mejorados los cuales cada uno es un probable proyecto de grado, entre ellos:

- Mapeo de la cadena de valor completa de producción y transformación de productos agrícolas.
- Logística de transporte de los productos agrícolas hacia ciudades intermedias y principales por medio de inteligencia artificial.
- Relación de producción de cadenas de valor con condiciones climáticas para estimación de recursos para suplir la demanda.

6.0.4. TESIS DE MAESTRÍA

Al analizar el área de implementación de la tesis es necesario desarrollar proyectos de base tecnológica que impacten en la cadena completa de valor del sector agrícola, para ello se presentan dos posibles proyectos con relevancia a implementar.

- Desarrollo de un proceso de estimación de estado actual de cadenas de valor por medio de Internet de las Cosas e Inteligencia Artificial.

Este proyecto se justifica en la medida que las instituciones de gobierno local de los 1.100 municipios del país no conocen el estado real de las cadenas de valor de productos insignias de su municipio. La información que se maneja la tienen las federaciones de productores independientes y no la comparten con los gobiernos, así pues al momento de crear políticas públicas las alcaldías y gobernaciones no saben donde ejecutar los recursos y quedan sometidas a lo que les digan las federaciones de productores.

Esta asimetría de información genera que los recursos no sean utilizados donde se necesita sino donde particulares asumen que se necesitan, con los riesgos de la mala adjudicación de los recursos bajo esta modalidad.

- Inteligencia artificial para el seguimiento a procesos jurídicos generando líneas jurisprudenciales.

Un caso puntual de la inteligencia artificial que vale la pena explorar es el análisis de textos jurídicos para generar líneas jurisprudenciales. El actuar de un juez esta determinado de las actuaciones que hayan tenido las altas cortes, este actuar parte de documentos donde consignan su foram de interpretar la ley. El proyecto toma como base estos documentos para evaluar los parámetros que llevan y crear una *línea jurisprudencial*, luego se puede crear una interfaz que verifique los accionares de los jueces regionales para identificar si siguen esta línea o están realizando sus fallos de forma distinta incurriendo en posibles hechos de corrupción.

BIBLIOGRAFÍA

- Alpaydm, E. (2010). Introduction to machine learning, second edition. *The MIT Press*, page 579.
- Analytics, I. (2016). Guide to iot solution development. *Gerencia tecnológica e informática*, page 53.
- Bagula, A. V. A. (2015). *Internet de las Cosas*. Latin Chapter of Internet Society.
- Barnaghi, P., Expert), U. W. I., Compton, M., of the SSN Ontology), E., Corcho, O., Castro, R. G., Graybeal, J., Herzog, A., Gesellschaft, F., Janowicz, K., Neuhaus, H., Nikolov, A., and Page, K. (2011). Semantic sensor network xg final report. *W3C Incubator Group Report*, page 30.
- Cambra, C., Sendra, S., Lloret, J., and Garcia, L. (2017). An iot service-oriented system for agriculture monitoring. *IEEE International Conference on Communications 2017 SAC Symposium Internet of Things Track*.
- Castillo, J. M. S. and Azkuenaga, X. Z. (2007). Valoración crítica de los modelos de gestión del conocimiento. *XXI Congreso Anual AEDEM, Universidad Rey Juan Carlos*, 2(1):50–65.
- Chavez-Ramirez, E., Gonzalez-Cervantes, G., Gonzalez-Barrios, J. L., Dzul-Lopez, E., Sanchez-Cohen, I., Lopez-Santos, A., and Chavez-Simental, J. A. (2013). Uso de estaciones climatológicas automáticas y modelos matemáticos para determinar la evapotranspiración. *Tecnología y Ciencias del Agua*, 4:115 – 126.
- Coates, R., Delwiche, M., Broad, A., and Holler, M. (2013). Wireless sensor network with irrigation valve control. *Computers and Electronics in Agriculture*, 99:13 – 22.
- Conventions, C. and Metadata (2018). Cf standard name table. <http://cfconventions.org/Data/cf-standard-names/50/build/cf-standard-name-table.html>, 50(1).
- de Acreditación Colombiana, O. N. (2007). La acreditación en colombia. *Asistencia Técnica al Comercio en Colombia*, 3(1):36.
- de la Información y las Comunicaciones, M. (2017). Centro de excelencia y apropiación en internet de las cosas. <http://www.cea-iot.org/>.
- de las Naciones Unidas para la Alimentación y la Agricultura, O. (2015). La fao y los 17 objetivos de desarrollo sostenible. *Desarrollo Sostenible*, page 75.

- Díaz, M. S. (2005). Breve inventario de los modelos para la gestión del conocimiento en las organizaciones. *Revista Cubana de Información en Ciencias de la Salud*, 13.
- Duarte, O. (1999). Sistemas de lógica difusa - fundamentos. *Revista Ingeniería e Investigación*, 3(42):22 – 30.
- Enciso, R. Z. (2011). *Cooplexity: Un Modelo de Colaboracion en Complejidad Para la Gestion en tiempos de Incertidumbre y Cambio*. Lulu Editors.
- FAO (2017). Evapotranspiración del cultivo. guías para la determinación de los requerimientos de agua de los cultivos. *Estudio FAO Riego y Drenaje*.
- Ferdoush, S. and Li, X. (2014). Wireless sensor network system design using raspberry pi and arduino for environmental monitoring applications. *The 9th International Conference on Future Networks and Communications (FNC-2014)*, 34:103 – 110.
- Group, O. C. A. W. (2016). Openfog architecture overview, white paper. Technical report, OpenFog.
- Gurovich, L. (2012). El suelo. In *Riego Superior Tecnificado*, pages 35–105. Ediciones Universidad Católica de Chile.
- Hargreaves, H. and Allen, G. (2003). History and evaluation of hargreaves evapotranspiration equation. *ascelibrary.org by University of Nebraska-Lincoln*.
- Haykin, S. (2009). Neural networks and learning machines third edition. *Pearson Education*, 3:1– 17.
- H.D., P. (2002). Intelligence, information technology and information. In *Intelligence, Information Technology and Information*, pages 51–74. Kuala Lumpur: Universidade Malaya.
- Hernández Sampieri, R., Fernández Collado, C., and Baptista Lucio, P. (2003). Metodología de la investigación. *La Habana: Editorial Félix Varela*, 2.
- Iglesias, H. J. P., González, L. L., Rodríguez, C. G., and Lema, S. S. (2012). Técnicas avanzadas de producción sostenible para agricultura de precisión en cooperativas. *CONAMA 2012 - Congreso Nacional del Medio Ambiente*, page 28.
- Jayaraman, P. P., Yavari, A., Georgakopoulos, D., Morshed, A., and Zaslavsky, A. (2016). Internet of things platform for smart farming: Experiences and lessons learnt. *Sensors*, 16(1884):1– 17.
- KaaIoT (2017). Iot implementation checklist for enterprise. Technical report, KaaIoT Technologies Inc.
- Kabilan and Selvi, S. (2016). Surveillance and steering of irrigation system in cloud using wireless sensor network and wi-fi module. *2016 Fifth Internation Conference on Recent Trends in Information Technology*.

- Khan, R., Khan, S. U., Zaheer, R., and Khan, S. (2012). Future internet: The internet of things architecture, possible applications and key challenges. *2012 10th International Conference on Frontiers of Information Technology (FIT): Proceedings*, pages 23 – 27.
- Khelifa, B., Ame, D., Amel, B., Mohamed, C., and Tarek, B. (2015). Smart irrigation using internet of things. *2015 4th International Conference on Future Generation Communication Technology*, pages 91 – 96.
- Kriesel, D. (2007). A brief introduction to neural networks. 1:244.
- Madanapalli, S. (2017). Standard for an architectural framework for real-time onsite operations facilitation (roof) for the internet of things, the roof computing. Technical report, IEEE Standard Association.
- Montoya, F., Gomez, J., Cama, A., Sierra, A. Z., Martinez, F., Cruz, J. L. D. L., and Agugliaro, F. M. (2013). A monitoring system for intensive agriculture based on mesh networks and the android system. *Computers and Electronics in Agriculture journal*, 99:14 – 20.
- Mora, J. P. and Arteaga, M. A. (2014). Metodología participativa de transferencia de tecnología en el cultivo del algodón en colombia. *Publicacoes Embrapa*, 1.
- Morcillo, C. G. (2014). Lógica difusa - una introducción práctica, técnicas de softcomputing. *Revista de la Escuela Superior de Informática*, 3:29.
- MQTT.ORG (2017). Mqtt version 3.1.1 plus errata 01 - oasis standard incorporating approved errata 01. <http://docs.oasis-open.org/mqtt/mqtt/v3.1.1/mqtt-v3.1.1.html>.
- Navarro-Hellin, H., del Rincon, J. M., Domingo-Miguel, R., Soto-Valles, F., and Torres-Sanchez, R. (2016). A decision support system for managing irrigation in agriculture. *Computers and Electronics in Agriculture*, 124:121 – 131.
- Navarro-Hellin, H., Torres-Sanchez, R., Soto-Valles, F., Albaladejo-Perez, C., Lopez-Riquelme, J., and Domingo-Miguel, R. (2014). A wireless sensors architecture for efficient irrigation water management. *Agricultural Water Management*, pages 64 – 74.
- Perez, V. A. and Urbaez, M. F. (2016). Modelos teóricos de gestión del conocimiento: descriptores, conceptualizaciones y enfoques. *Entreciencias, dialogos en la sociedad del conocimiento*, 4:202 – 227.
- Poole, D., Mackworth, A., and Goebel, R. (2006). Computational intelligence: A logical approach. *Oxford University Books*, page 500.
- Putjaika, N., Phusae, S., Chen-Im, A., Phunchongharn, P., and Akkarajitsakup, D. (2016). A control system in an intelligent farming by using arduino technology. *Proceedings of the 2016 5th ICT International Student Project Conference, ICT-ISPC 2016*, 16:59 – 63.
- Quintero, J. F. L., Giraldo, G. A. M., and Bernal, H. D. (2014). Gestión del conocimiento en riesgos de desastres. *12th Latin American and Caribbean Conference for Engineering and Technology*.

- Rajalakshmi and Mahalakshmi, D. (2016). Iot based crop-field monitoring and irrigation automation. *Proceedings of the 10th International Conference on Intelligent Systems and Control, ISCO 2016*.
- Ray, P. P. (2017). Internet of things for smart agriculture: Technologies, practices and future direction. *Journal of Ambient Intelligence and Smart Environments*, 9:395 – 420.
- Richard G. Allen, Luis S. Pereira, D. R. (1998). Crop evapotranspiration - guidelines for computing crop water requirements - fao irrigation and drainage paper 56. *FAO - Food and Agriculture Organization of the United Nations*.
- Sales, N., Remedios, O., and Arsenio, A. (2015). Wireless sensor and actuator system for smart irrigation on the cloud. *IEEE World Forum on Internet of Things, WF-IoT 2015 - Proceedings*, 15.
- Smarsly, K. (2013). Agricultural ecosystem monitoring based on autonomous sensor systems. *2nd International Conference on Agro-Geoinformatics: Information for Sustainable Agriculture, Agro-Geoinformatics*, 1.
- Suryakant, S., Surya, D., and Jagarlapudi, A. (2017). Interoperable agro-meteorological observation and analysis platform for precision agriculture: A case study in citrus crop water requirement estimation. *Computers and Electronics in Agriculture*, 138:175 – 187.
- Torres, J. H. (2017). Inteligencia integral de negocios. *Doctorado en ingeniería. Universidad Distrital Francisco José de Caldas*.
- Vanegas, L. P.P. (2015). *Mecanismos de los modelos de gestión del conocimiento orientados a la generación de ideas innovadoras en Cemex Colombia*. Universidad Nacional de Colombia.
- W3C-Incubator-Group (2011). Semantic sensor network xg final report. <https://www.w3.org/2005/Incubator/ssn/XGR-ssn-20110628/>.
- Wesney, E. (2008). Building a competitor intelligence organization: adding value in a information function. In *Building a competitor intelligence organization: adding value in a information function.*, pages 25–35. Oxford University Press.
- Yu, X., Wu, P., Han, W., and Zhang, Z. (2013). A survey on wireless sensor network infrastructure for agriculture xiaoqing. *Computer Standards and Interfaces*, 35:59 – 64.
- Zanella, A., Bui, N., Castellani, A., Vangelista, L., and Zorzi, M. (2014). Internet of things for smart cities. *IEEE Internet of Things Journal*, 2(1):50–65.
- Zhang, X., Zhang, J., Li, L., Zhang, Y., and Yang, G. (2017). Monitoring citrus soil moisture and nutrients using an iot based system. *Sensors*, 17:1 – 10.
- Zhou, H. (2013). *The Internet of Things in the Cloud, a Middleware Perspective*. CRC Press.



ANEXO: PRODUCCIÓN BIBLIOGRÁFICA

En este Anexo se presentan apartes del artículo presentado a la revista *Visión Electrónica* el cual ya fue aprobado para publicación. Se presenta el título, resumen, palabras claves e información adicional tanto en español como en inglés el resumen.

A.0.1. TÍTULO

Smart irrigation: Data capture process based on knowledge management / Riego inteligente: proceso de captura de datos basado en gestión del conocimiento

A.0.2. RESUMEN

Abstract: This paper presents the process of acquiring environmental data that feed an intelligent irrigation control system, which based on the calculation of the evapotranspiration of a crop manages to calculate the water needs of the crop to supply them. It presents the problem of irrigation because a solution based on the Internet of Things (IoT) is considered satisfactory, specifying the variables involved in the process and the characteristics of the data produced by the sensors. After this, it develops the process of capturing data on an IoT architecture based on knowledge management and with the sensing, communication, and analytical phases, referring to the R software components that have been developed to carry out this process, culminating with the projections of irrigation analytics. As irrigation is the main aspect of crop yield, a need inherent to the field sector that is not yet automated and that seeks solutions to the conditions of the Colombian countryside is supplied. Keywords: Evapotranspiration, FAO, Irrigation, MQTT.

Resumen: El presente artículo describe el proceso de adquisición de datos medioambientales que alimentan un sistema de control de riego inteligente el cual, basado en el cálculo de la evapotranspiración de un cultivo, logra calcular las necesidades hídricas del mismo para suplirlas. Luego de plantearse la problemática del riego, y la justificación de una solución basada en Internet de las Cosas (IoT) como satisfactoria, se precisan las variables que intervienen en el proceso y las características de los datos que producen

los sensores; se desarrolla el proceso de captura de datos sobre una arquitectura IoT basada en gestión del conocimiento con las fases de: sensado, comunicación y analítica, refiriendo los componentes del software R que se han implementado para realizar este proceso, culminando con las proyecciones de analítica del riego. Al ser el riego el aspecto principal del rendimiento de un cultivo se concluye que se suple una necesidad inherente al sector del campo -que aún no está automatizado- proponiéndose una solución para las condiciones específicas del campo colombiano.

A.0.3. PALABRAS CLAVE

Keywords: Evapotranspiración, FAO, Irrigación, MQTT.

A.0.4. DETALLES

Preparación de Artículos revista VISIÓN ELECTRÓNICA: algo más que un estado sólido Fecha de envío: 23-09-2018 Fecha de recepción: 24-09-2018 Fecha de aceptación: 12-11-2018