

cujae

Prácticas Profesionales II

Sistema para el entrenamiento de operarios en la Industria
Alimentaria Cubana

Autor: Mónica Montoto Montané

Tutor: Dra. Raisa Socorro Llanes

La Habana, Cuba 4 de julio de 2022

Resumen

La Industria Alimentaria en Cuba juega un rol muy importante dentro de la economía, por lo que debe evitarse a toda costa un problema en este sector. Actualmente presenta la necesidad de capacitar a todos los operarios de sus plantas, ante los procesos productivos que en ellas se desarrollan, pues existen muchas irregularidades a la hora de operar estos procesos. Cuando existe una falla, no siempre están presentes los expertos para su posible solución, ya que no todos los técnicos están preparados para tomar decisiones correctas en determinadas circunstancias. Es por esto, que al ocurrir una falla, se detiene el proceso productivo hasta que aparezca el experto, provocando pérdidas en numerosas áreas y afectando la economía. A raíz de esta situación el Instituto de Investigaciones de la Industria Alimentaria (IIIA), en conjunto con la facultad de Ingeniería Informática e Ingeniería Química de la CUJAE, desarrolló un Sistema Experto para el Control de Procesos Químicos (SECPROIT) y un Generador de Bases de Conocimientos.

El SECPROIT tiene como objetivo evaluar las decisiones que toman los operarios ante diferentes situaciones críticas y brindar al jefe de área los resultados de las evaluaciones y como se van superando los mismos ante cada proceso. En estos momentos, el sistema presenta ciertas limitaciones que impiden su correcto funcionamiento, por lo que es necesario realizar una actualización del mismo.

Palabras claves: Industria Alimentaria, proceso productivo, operario, Instituto de Investigaciones de la Industria Alimentaria (IIIA), Sistema Experto para el Control de Procesos Químicos (SECPROIT), Generador de Bases de Conocimientos.

Índice general

Introducción						
1.	Fundamentos Teóricos					
	1.1.	Sistem	nas de Capacitación Automatizados	6		
		1.1.1.	Características de un Sistema de Capacitación	7		
		1.1.2.	Importancia de los Software de Capacitación	7		
		1.1.3.	Fases de un Proceso de Evaluación de Conocimiento	7		
		1.1.4.	Tipos de Preguntas en un Sistema de Capacitación Automatizado	8		
		1.1.5.	¿Cómo verificar la validez de las respuestas?	9		
		1.1.6.	¿Cómo se evalúa una Capacitación?	10		
	1.2.	Sistem	nas Expertos	10		
		1.2.1.	Componentes de un Sistema Experto	10		
		1.2.2.	Ventajas de un Sistema Experto	11		
		1.2.3.	Desventajas de un Sistema Experto	12		
		1.2.4.	$\ensuremath{\partial}$ Por qué usar un Sistema Experto en este Entrenador?	12		
	1.3.	Sistem	na de Entrenamiento SECPROIT	13		
		1.3.1.	¿Cómo funciona el SECPROIT?	14		
		1.3.2.	Limitaciones Actuales del Sistema	14		
		1.3.3.	Nuevas Acciones Sugeridas	15		
	1.4.	Conclu	usiones Parciales	16		
2.	Nue	evo Sis	etema de Entrenamiento	17		
	2.1.		lado del Negocio	17		
		2.1.1.	Modelo de Dominio	17		
		2.1.2.	Reglas del Negocio	18		
		2.1.3.	Diagrama de Actividades	19		
	2.2.		ra de Requisitos	19		
		2.2.1.	Requisitos Funcionales	20		
		2.2.2.	Requisitos No Funcionales	20		
	2.3.		de Uso del Nuevo Sistema	$\frac{-3}{21}$		
			Actores del Sistema	22		

Referencias bibliográficas								
	2.4.1.	Diagrama de Entidades y Relacio	ones (DER)		23			
2.4.	Model	de Datos			23			
	2.3.2.	Diagrama de Casos de Uso del N	uevo Sistema		22			

Índice de tablas

Índice de figuras

1.1.	Etapas para garantizar un buen Sistema de Capacitación	8
1.2.	Componentes mínimos requeridos en un Sistema Experto	11
1.3.	Estructura de un archivo drl	14
2.1.	Modelo de Dominio del Nuevo Sistema	18
2.2.	Diagrama de Actividades del Nuevo Sistema	19
2.3.	Diagrama de Casos de Uso del Nuevo Sistema	23

Introducción

En la actualidad, el desarrollo tecnológico se ha convertido prácticamente en una obligación para la mayoría de los entornos laborales, industriales y sociales. En un mundo tan globalizado, que avanza y se supera con una rapidez vertiginosa, no resulta de extrañar que se exijan mejores condiciones y prestaciones. A partir de estas demandas, surge la necesidad de establecer nuevas técnicas en las grandes empresas que, de manera automática, faciliten todo tipo de actividades, desde lo laboral y comercial, hasta lo cotidiano [1]. Estas exigencias requieren nuevos retos, además de los tecnológicos. Los avances técnicos traen consigo la necesidad de contar con factores humanos capacitados para responder ante el nuevo desarrollo científico [2]. Estadísticamente, la tarea no resulta sencilla para aquellos trabajadores que su vida ha sido lejos de un ordenador.

Cuba no está ajena a esta nueva revolución tecnológica. El país está abocado a la creación y fortalecimiento de programas de posgrado que consoliden, complementen, profundicen y originen nuevos conocimientos, inmersos en áreas relacionadas con estos cambios de la era digital. De esta forma se proporcionan herramientas para afrontar retos, tomar decisiones y solucionar problemas de la sociedad cubana en general [1]. El empleo de una tecnología de software propia reduce el tiempo de desarrollo y puesta marcha de la misma, minimiza el costo, aporta gran flexibilidad y eficiencia en las operaciones, mejora significativamente la trazabilidad de los procesos y facilita el trabajo a partir de diseños amigables de la interfaz hombre-máquina [3]. Es por ello que, nuestro país, constantemente trabaja en función de desarrollar nuevos sistemas que permitan un avance técnico.

Por citar un ejemplo, la Industria Farmacéutica Cubana está compuesta por procesos tecnológicos de elevada complejidad, que necesitan cumplir con exigentes indicadores de calidad en la producción de medicamentos. A esta industria se le empleó una metodología que permitió automatizar la mayoría de sus procesos. Esta, ha significado un ahorro al país superior a tres millones y medio en moneda convertible en los últimos cinco años, un gran beneficio gracias al avance tecnológico [4].

En ocasiones, las ventajas competitivas que se obtienen con el uso de la tecnología, pueden verse afectadas por diversos riesgos, lo que provoca una posibilidad de pérdida hacia el futuro. Aunque son un hecho probabilístico sobre el que se tiene cierta incertidumbre, porque pueden ser provocados por factores internos o externos a la or-

ganización, tienen el poder de originar efectos indeseados en el cumplimiento de los objetivos de las instituciones. Ante esta realidad, ha surgido como necesidad la gestión de riesgos, con el propósito de lograr ejercicios de planeación y ejecución más acertados. De esta situación no escapa la Industria Cubana y como parte de ella la Industria Alimentaria [5].

La Industria Alimentaria Cubana juega un rol muy importante dentro de la economía del país. Se caracteriza por trabajar de manera ininterrumpida, con un personal que cambia frecuentemente. Esta situación dificulta el entrenamiento de sus procesos productivos, ya que no cuenta con suficientes especialistas para la capacitación de los trabajadores. Al generarse una falla en un proceso, existe el inconveniente de una posible ausencia de expertos que encuentren una solución al problema, ya que todos los técnicos no tienen el mismo nivel de preparación para tomar decisiones ante fallas en determinadas circunstancias. En este caso, resulta entonces necesario paralizar el proceso productivo hasta localizar al experto que solucione el problema detectado, lo que genera pérdidas de recursos e incumplimiento del tiempo de producción [6].

El Instituto de Investigación de la Industria Alimentaria (IIIA) posee un capital humano especializado y de prestigio que, a lo largo de los años, logró resultados científicotécnicos consistentes en la creación de aditivos, extensores alimenticios y la fortificación con vitaminas y minerales de alimentos seleccionados. El propósito de esta institución es lograr ventajas competitivas sostenibles en las industrias de nuestro país anaMailen. Para darle solución al problema de la Industria Alimentaria Cubana, este instituto, en conjunto con las facultades de Ingeniería Química e Ingeniería Informática (ambas de la CUJAE), desarrolló un Sistema Experto para el Control de Procesos Químicos (SEC-PROIT) y un Generador de Bases de Conocimientos. Estos sistemas fueron diseñados para capacitar a los operadores de las plantas de producción ante las diferentes fallas que se puedan presentar, solucionando la problemática inicial planteada anteriormente.

El Generador de Bases de Conocimientos es el encargado de crear las bases de información para los diferentes procesos químicos que ocurren en una fábrica. Cada base contiene un grupo de variables, cada una con sus características y clasificaciones. De cada variable se conocen las causas que pueden provocar un estado de riesgo o alarma en las mismas. De cada causa, se conocen las recomendaciones a seguir para minimizar los riesgos en el proceso productivo. A partir de estos datos, el generador elabora un conjunto de ficheros que luego serán utilizados por el SECPROIT [7].

El Sistema Experto para el Control de Procesos Químicos (SECPROIT) tiene como objetivo asesorar la toma de decisiones durante la ejecución de un proceso. A partir de los datos obtenidos por el Generador de Bases de Conocimiento, este sistema produce un conjunto de entrenamientos para los operarios de la fábrica. En cada prueba, el usuario debe escoger las variables que considere en estado de alarma, señalar sus posibles causas y seleccionar las recomendaciones a seguir. Cada prueba contiene tres etapas:

variables, causas y recomendaciones. Estas evaluaciones son utilizadas para advertir los conocimientos de los trabajadores y la capacidad que posee cada uno [8].

Actualmente el sistema posee ciertas limitaciones que afectan el resultado final buscado:

- el operario solo puede realizar el entrenamiento una vez, es decir, solo se evalúa el contenido de un proceso una vez, lo que limita la capacidad de superación del trabajador
- todas las etapas de evaluación se encuentran continúas y no influyen de manera determinante en el resultado final, es decir, se puede suspender una etapa y aprobar dos, sin importar la relevancia de la primera
- la última etapa (las recomendaciones) no se evalúan correctamente, ni muestran los resultados obtenidos
- la interfaz de usuario no es intuitiva, por lo que algunos usuarios no llegan a comprender su funcionamiento
- las interfaces del sistema son poco atractivas con colores muy oscuros, lo que provoca que el usuario niegue su uso

Partiendo de esta **situación problemática** resulta necesario desarrollar un nuevo sistema, tomando como base el SECPROIT. Esta actualización debe poder resolver el **problema de investigación** principal: ¿Cómo lograr una correcta y completa capacitación de los operarios ante los procesos productivos de la fábrica?

Es por ello que, como **objetivo general** de la investigación, se tiene:

 Generar un nuevo sistema que no contenga las limitaciones del SECPROIT, pero que realice las mismas funciones que este.

Para poder conseguir este objetivo general, se deben cumplir los siguientes **objeti-** vos específicos:

- Analizar el sistema SECPROIT para definir las operaciones que se pueden aprovechar y las limitaciones que se deben eliminar
- Modelar el nuevo sistema e identificar los requisitos que debe cumplir
- Diseñar la base de datos del nuevo sistema a partir del modelado anteriormente realizado
- Completar el diseño de la solución a partir de los resultados obtenidos en los objetivos anteriores

- Programar el nuevo sistema
- Realizar diseño y ejecución de pruebas al nuevo sistema creado
- Desplegar la nueva aplicación en la facultad de Ingeniería Química

Con estos objetivos se deben cumplir un conjunto de tareas que facilitarán el proceso de desarrollo del sistema. Estas **tareas específicas** son:

- Investigar los tipos de sistemas de entrenamientos existentes, así como ejemplos de evaluaciones y calificaciones que se pueden aplicar en los mismos
- Realizar diagrama de casos de uso, diagrama de base de datos, diagrama de clases y algún otro tipo de diagrama necesario para un mejor entendimiento de los diseños propuestos
- Implementar la carga de la base de conocimiento, que permite realizar los entrenamientos en el nuevo sistema
- Implementar generación de preguntas asociadas a la evaluación del estado de las variables
- Implementar generación de preguntas asociadas a la evaluación de las causas asociadas a las causas del estado de una variable
- Implementar generación de preguntas asociadas a la evaluación de las recomendaciones asociadas al estado de una variable
- Implementar la configuración del proceso de evaluación asociado a un proceso específico
- Implementar el proceso de conexión con el JBOSS para obtener las respuestas correctas para las preguntas asociadas a las causas y recomendaciones
- Implementar el proceso de evaluación parcial de cada etapa
- Implementar la funcionalidad que permita obtener los reportes asociados a los avances de cada operario, proceso y área
- Implementar el proceso de evaluación integral del operario en un proceso
- Implementar la funcionalidad que permita gestionar la administración del sistema
- Ejecución de las pruebas y experimentos al nuevo sistema

Los **objetos de estudio** de esta investigación son los sistemas de entrenamiento, los sistemas expertos y los sistemas de información inteligentes basados en reglas de producción. De este conjunto, se centra la atención en el **campo de acción** que comprende el Sistema Experto para el Control de Procesos Químicos (SECPROIT).

Como artefacto de salida se logrará un sistema de entrenamiento (una aplicación práctica) que permite capacitar a los operarios de las Industrias Alimentarias. En una primera prueba, se desea implementar este sistema resultante en las Industrias Azucareras Cubanas, obteniendo un valor práctico de gran importancia para la sociedad.

Capítulo 1

Fundamentos Teóricos

El presente capítulo abarca los principales temas que se abordan a lo largo de la investigación. Se detalla qué es un sistema de capacitación automatizado, su importancia, sus tipos de preguntas y los modelos de calificación que siguen. Además, se plantea el concepto de sistemas expertos, junto a sus características más notables y su importancia. También estarán explicadas algunas de las funcionalidades del Sistema de Entrenamiento SECPROIT, sus componentes y rendimiento. Al final del capítulo se observan conclusiones parciales a modo de resumen del mismo.

1.1. Sistemas de Capacitación Automatizados

La capacitación laboral es un método aplicado por las empresas para que su personal adquiera nuevos conocimientos profesionales. Por lo general, se produce ante un ascenso o incorporación, aunque no son los únicos motivos. Lo ideal es que se desarrolle de forma continúa, ya que la constante formación del personal deriva en resultados positivos tanto para el grupo de trabajo como para la organización en la que se realiza [9]. Surge en el mundo como respuesta a la necesidad de mejorar permanentemente la calidad y formación de recursos humanos.

Basándose en el concepto anterior, un sistema de capacitación automatizado, también conocido como sistema de entrenamiento, es un método de enseñanza alternativo creado para el adiestramiento de los trabajadores. Se basa, principalmente, en un software que permite el aprendizaje de los usuarios sin necesidad de una supervisión constante. Generalmente, resulta más efectivo que las prácticas de enseñanza presencial, debido a que el estudiante trabaja solo y puede determinar su propia velocidad de aprendizaje, usando una amplia variedad de herramientas y métodos para la transferencia del conocimiento [10].

A modo de resumen, es un software que brinda una solución de recursos humanos, ayuda en la formación de los trabajadores y aumenta la productividad empresarial.

1.1.1. Características de un Sistema de Capacitación

Un sistema de capacitación, ya sea automatizado o no, ofrece diferentes aplicaciones en función del modelo de negocio que utiliza. Su versatilidad permite adaptarse a las necesidades particulares de cada sector. Sin embargo, según [11], la mayoría de los sistemas contienen las mismas características:

- Son capaces de gestionar los distintos cursos impartidos, la asistencia y la inversión en formación de la empresa.
- Asignan a los empleados que deberán asistir y los profesionales responsables de analizar sus resultados.
- Detectan las carencias formativas del personal antes de que influyan en el desarrollo del trabajo.
- Clasifican las distintas actividades formativas en base a su categoría y catálogo.
- Registran y consultan el progreso del aprendizaje de los empleados en tiempo real.

1.1.2. Importancia de los Software de Capacitación

Un sistema de entrenamiento asistido por computadora permite ofrecer el mismo nivel de adiestramiento para cada usuario del sistema, en cuanto a rigor y evaluación. Uno de los problemas de la capacitación de los empleados de manera presencial, es que las sesiones son frecuentemente inconsistentes, y las diferencias en el nivel de habilidad del formador pueden tener un impacto significativo en el éxito del empleado. Al contar con un sistema automatizado, solo se necesita una base de conocimientos para garantizar el mismo nivel de entrenamiento para todos los capacitados [12].

1.1.3. Fases de un Proceso de Evaluación de Conocimiento

Un proceso de evaluación de conocimiento, debe estar integrado por cinco etapas (Figura 1.1), asegura [13]. Cada una de ellas, va a marcar un conjunto de acciones, que al final se interpretarán como un buen entrenamiento:

Recogida de Datos: es la recopilación sistemática de toda la información a lo largo del proceso completo de enseñanza-aprendizaje. Debe tener concordancia con los objetivos, ser suficiente, representativa, relevante y ponderada, en función del peso otorgado a cada objetivo. En los sistemas en línea estas posibilidades de registrar evidencias son inmensas.

- Puntuación de las Pruebas: se realiza una vez medidos, de manera cuantitativa o cualitativa, los distintos bloques de información, con las ponderaciones, criterios e indicadores que se hayan establecido.
- Juicio de Valor: puede hacerse limitándose a criterios de grupo (evaluación normativa), refiriéndose a criterios de superación de objetivos y/o contenidos (evaluación de criterio), o teniendo en cuenta la personalidad, posibilidades y limitaciones del propio sujeto del aprendizaje (evaluación personalizada).
- Toma de Decisiones: habitualmente denominada calificación, trae consigo una serie de consecuencias personales, administrativas, económicas y laborales. Se basa en la decisión a partir del resultado. La acción resultante influye directamente en el adiestrado. Ejemplos de decisiones son: selección/exclusión, promoción/recuperación/repetición, rebajar a niveles anteriores, ampliación, certificación/reprobación, entre otras.
- Información a los Interesados: es la etapa final, que ha de llegar a diferentes destinatarios, aunque principalmente y de forma adecuada, a los capacitados. Es la confirmación de que concluye el entrenamiento.



Figura 1.1: Etapas para garantizar un buen Sistema de Capacitación.

1.1.4. Tipos de Preguntas en un Sistema de Capacitación Automatizado

A medida que avanza el tiempo, se generan nuevos métodos de estudio, y con estos, nuevas formas de preguntar y calificar. Sin embargo, a la hora de diseñar un sistema automatizado, no es menos cierto que existen algunas variantes más sencillas y por ende, más utilizadas. Según [14] los tipos de preguntas que mayormente se emplean en un sistema de este tipo son:

- Verdadero o Falso: contienen una declaración que se debe indicar si es verdadera
 o no. Permiten responder en poco tiempo, son fáciles, rápidos de calificar y se
 corrigen de forma automática.
- Opción Múltiple: se componen de una pregunta (raíz) con múltiples respuestas posibles. Pueden incluir múltiples opciones válidas, en cuyo caso, podrían darse por superada al marcar cualquiera de ellas o cuando se marquen todas. Se caracterizan por ser fáciles y rápidas de calificar, se corrigen automáticamente y se pueden utilizar para evaluar los conocimientos en una amplia gama de contenidos.
- Emparejar, Relacionar u Ordenar: por lo general se emparejan cada una de las opciones del primer bloque con las opciones dadas en el segundo bloque, o se ordenan bloques de modo que quede una secuencia correcta de acuerdo a un patrón previamente establecido. Se suelen usar en aquellos cursos donde la adquisición de conocimientos muy detallados es un objetivo importante. Son preguntas fáciles de diseñar, rápidas de calificar y se corrigen automáticamente. Estadísticamente, se tarda más en responder que las preguntas anteriores.
- Respuesta Corta: basta con que se escriban un par de palabras o una frase sencilla. Una alternativa más común a este tipo de preguntas es la de cubrir espacios en blanco con una palabra. Este tipo de preguntas son muy útiles para que se demuestren los conocimientos basados en hechos o palabras claves. La dificultad para calificarlas depende del estilo que se decida emplear.

1.1.5. ¿Cómo verificar la validez de las respuestas?

Una vez terminado el entrenamiento, mediante el proceso de evaluación, se debe comprobar cuáles de los resultados obtenidos son correctos y cuáles no. Para ello se deben comparar las respuestas del evaluado con una fuente de confianza, que contenga la información verídica de lo que se está tratando. Estas fuentes de confianza se conocen por el nombre: bases de conocimiento. A partir de ellas, se verifica si los datos en las respuestas del evaluado coinciden con la información real contenida en la base. Este proceso puede realizarse tanto de manera manual, semi-automática o automática [15].

Al tratarse de un sistema de capacitación automatizado, por lo general, el método utilizado para validar las respuestas es el automático. De esta forma se facilita el trabajo para aquellos que deben evaluar a un personal abundante. Según [16], la manera más efectiva y eficiente de evaluar estos sistemas es a partir del uso de un Sistema Experto.

1.1.6. ¿Cómo se evalúa una Capacitación?

La evaluación y la calificación de una prueba, no pueden depender de un solo instrumento o técnica, ya que de esa forma solo se mide un tipo de aprendizaje. Los criterios para calificar que se designen, serán los porcentajes de valor que se establezcan a cada resultado de las actividades realizadas y a su resultado final. Se debe tomar en cuenta tanto la exactitud de la respuesta, como el proceso que se siguió para llegar a la misma, así como la cantidad de intentos necesarios utilizados para hallar la solución correcta. Una evaluación posee dos objetivos principales: analizar en qué medida se han cumplido los objetivos, para detectar posibles fallas en el proceso y poder superarlas, y proporcionar la reflexión de los que realizaron el entrenamiento en torno a su propio proceso de aprendizaje (metacognición) [17].

A modo de resumen, para realizar una correcta evaluación se deben tener en cuenta tantas herramientas como parámetros influyan. En este caso en particular: la puntuación de las respuestas y la cantidad de intentos.

1.2. Sistemas Expertos

Los sistemas expertos resuelven problemas que normalmente son solucionados por expertos humanos. Para resolverlos, necesitan acceder a una importante base de conocimiento sobre el dominio, que debe construirse de la manera más eficiente posible. Utilizan uno o más mecanismos de razonamiento, para aplicar este conocimiento a los problemas que se le proponen. Cuentan con un mecanismo para explicar a los usuarios que han confiado en ellos, lo que han hecho [16].

Una forma de contemplar los sistemas expertos es que representan la mayor parte de la Inteligencia Artificial (IA) aplicada. Un sistema experto en IA se define como un programa informático que tiene la capacidad de representar y razonar sobre el conocimiento [18].

1.2.1. Componentes de un Sistema Experto

En [21] se comentan los diferentes componentes que integran un sistema experto. Aunque pueden contar con un número mayor de componentes, los mínimos requeridos son (Figura 1.2):

- Motor de Inferencia: es el corazón del sistema experto. Su cometido principal es sacar conclusiones aplicando el conocimiento a los datos. Estas conclusiones pueden estar basadas en conocimiento determinista o conocimiento probabilístico.
- Base de Conocimientos: consiste en un conjunto de objetos y un conjunto de reglas, que gobiernan las relaciones entre esos objetos. La información que almacena

es de naturaleza permanente y estática, es decir, no cambia de una aplicación a otra. Se debe diferenciar entre los datos y el conocimiento. El conocimiento se refiere a afirmaciones de validez general tales como reglas, distribuciones de probabilidad, entre otras. Los datos se refieren a información relacionada con una aplicación en particular [20].

- Mecanismo de Aprendizaje: controla el flujo del nuevo conocimiento que va del experto humano a la base de conocimiento. El sistema determina qué nuevo conocimiento se necesita, o si el conocimiento es realidad, es decir, si debe incluirse, y en caso necesario incorporar dicho conocimiento.
- Interfaz de Usuario: se constituye como la interfaz entre el sistema experto y el usuario. Para que sea efectiva debe incorporar mecanismos para mostrar y obtener información de forma sencilla y agradable.



Figura 1.2: Componentes mínimos requeridos en un Sistema Experto.

1.2.2. Ventajas de un Sistema Experto

El uso de un sistema experto en cualquier ámbito social, resulta favorable de diversas maneras. En [19] se mencionan algunas de estas ventajas:

- Mientras que un experto humano tiene limitaciones y percances propias de su condición humana, el Sistema Experto no sufre de estas cuestiones y se convierte en una herramienta estable y fiable para su entorno.
- Sus actividades son completamente replicables y siempre contesta de la misma manera a menos que se le cambie el diseño.
- La velocidad de procesamiento es mayor al de un ser humano.
- Debido a la escasez de expertos humanos en determinadas áreas, los sistemas expertos pueden almacenar su conocimiento para cuando sea necesario aplicarlo.

- Pueden ser utilizados por personas no especializadas para resolver problemas.
- Pueden ser usados como sistema de aprendizaje.
- Si se evalúa el costo total del empleo de esta tecnología, la replicabilidad y estabilidad, asociado a la seguridad que provee, resulta una ecuación favorable, aun considerando que las inversiones iniciales pueden ser relativamente elevadas.

1.2.3. Desventajas de un Sistema Experto

A su vez, el empleo de estos sistemas puede traer consigo alguna desventaja. Nuevamente, en [19] se mencionan algunas de estas:

- Para actualizarlos se necesita de reprogramación, siendo una de sus limitaciones más acentuadas.
- Poseen elevado costo en dinero y tiempo, además que estos programas son poco flexibles a cambios y de difícil acceso a información no estructurada.
- Carecen de sentido común, para un sistema de este tipo no hay nada obvio.
- No podemos mantener una conversación informal con estos sistemas.
- Para un sistema experto es muy complicado aprender de sus errores o de errores ajenos.
- No son capaces de distinguir cuales son las cuestiones relevantes de un problema y separarlas de cuestiones secundarias.

Sin embargo, estos problemas no solo los presentan los sistemas expertos, puesto que la inteligencia artificial aún no ha podido desarrollar sistemas que sean capaces de resolver problemas de manera general o de aplicar el sentido común para resolver situaciones complejas o controlar situaciones ambiguas. Es por eso que, a pesar de sus desventajas, los sistemas expertos son considerados una gran ayuda y un enorme avance, en especial, en los sistemas de capacitación.

1.2.4. ¿Por qué usar un Sistema Experto en este Entrenador?

La Industria Alimentaria Cubana trabaja de manera ininterrumpida, con un personal laboral que cambia frecuentemente. Es por ello que no cuentan con la cantidad de expertos necesarios para que se atiendan, de manera constante, todos los procesos que en ella ocurren. La capacitación presencial de los trabajadores no es una opción válida, puesto que sufren de constante cambio de personal y los expertos no poseen

el tiempo para capacitarlos a todos. Por estas razones, se hace imprescindible la tarea de entrenar, de alguna manera, a los operarios de esta industria, ya que lográndolo, se garantiza una mejor respuesta ante una falla en algún proceso productivo [6].

Un método muy eficiente para entrenar a este personal, es mediante un sistema de capacitación digital, porque, como bien se fundamentó en la sección anterior, así pueden aprender por sí solos, sin la presencia de un experto y a su propio ritmo, y pueden ser evaluados de las mismas maneras, sin distinciones de preferencias o errores humanos. Con una capacitación digital también se logra hacer más sencillo el trabajo del jefe de área, ya que solo deberá seleccionar al mejor entre un conjunto de resultados, sin tener que aplicar pruebas o entrevistas.

Los procesos que se manejan en esta industria se caracterizan por tener datos estructurados en variables, donde cada variable tiene sus características principales, las causas que pueden volverla inestable y las recomendaciones a seguir para que no ocurra un incidente o qué hacer en caso de que ya ocurriese. Este grupo de datos conforma una base de conocimientos.

Si se analiza en conjunto los datos que se tienen y las necesidades que se desean cumplir, se llega a la conclusión de que se quiere crear un sistema de entrenamiento, para capacitar a un conjunto de trabajadores debido a la ausencia de expertos humanos, y para ello se cuenta con una base de conocimientos y un mecanismo de aprendizaje diseñado. A modo de resumen, un sistema experto.

1.3. Sistema de Entrenamiento SECPROIT

Para brindar una solución a la principal problemática de esta investigación, en el año 2018 se crearon dos aplicaciones que trabajarían de manera conjunta. La primera es un Gestor de Bases de Conocimientos que permite crear las bases de información que se necesitan para evaluar a los trabajadores. La segunda es conocida como Sistema Experto para el Control de Procesos Químicos (SECPROIT), que es un sistema diseñado para capacitar a los operarios de las fábricas. Ambos sistemas fueron desarrollados en la Universidad Tecnológica de La Habana José Antonio Echeverría (CUJAE), entre las facultades de Ingeniería Química e Ingeniería Informática. Ambas aplicaciones sirvieron para defender el título de Ingeniería.

Aunque estas aplicaciones se encuentren estrechamente vinculadas a este trabajo, en la mismo se abordará más, el sistema SECPROIT. El objetivo principal es lograr un sistema semejante con mejores prestaciones y funcionalidades, una actualización.

1.3.1. ¿Cómo funciona el SECPROIT?

En el sistema existen tres roles fundamentales: el administrador que se encarga de llenar los datos de los usuarios, el especialista que inserta los procesos para los entrenamientos y supervisa los resultados obtenidos por los trabajadores de su área laboral, y el operario que realiza la capacitación. Cada usuario tiene un nombre y una contraseña, y con cada rol aparecen funcionalidades únicas. Un rol no puede realizar las acciones de otro.

En el caso de los especialistas, deben registrar de cada proceso el nombre, una imagen si la posee, un fichero tipo anm y un archivo drl. Los ficheros anm contienen los nombres de todas las variables que influyen en el proceso, las características, todas las causas que pueden afectar el proceso y todas las recomendaciones para todas las causas descritas. El archivo drl contiene las reglas que enlazan las variables con sus causas y las causas con sus recomendaciones. Dichas reglas siguen una estructura específica: comienza con un patrón para el nombre de la regla, luego presenta los posibles atributos que posee, la sentencia que se debe cumplir y las acciones que realizar si se cumple la sentencia (Figura 1.3). Este archivo es lo que se conoce como Motor de Reglas (Drools). Esta información es extraída del sistema Gestor de Bases de Conocimientos [7].

Figura 1.3: Estructura de un archivo drl.

Por su parte, el operario entrará a realizar aquellos entrenamientos que el especialista haya validado. Una vez comenzado, deberá señalar, de un grupo de variables, las que por su estado estén en peligro de inestabilidad. Si ha seleccionado correctamente pasa a la siguiente etapa, donde debe escoger qué causa el estado de las variables que prefirió. Por último deberá seleccionar qué recomendaciones seguir para cada causa que señaló. Este entrenamiento es un proceso continuo donde, si decides abandonar sin haber terminado, no se guardarán los cambios y aparecerá como prueba no superada. La puntuación está dada de manera genral, como porcentaje entre las notas de las tres etapas [8].

1.3.2. Limitaciones Actuales del Sistema

Actualmente el sistema posee ciertas limitaciones que afectan la experiencia de usuario, sus principales funcionalidades y en general, el disfrute de la aplicación:

- Los entrenamientos de un proceso son únicos, es decir, solo se evalúa ese contenido una vez, sin brindar la posibilidad de mejorar o aprender.
- Las etapas del entrenamiento son continuas, sin posibilidad de pausa o descanso.
- Si no se aprueba una de las etapas del entrenamiento, el operario deberá continuar y concluirlo, aunque ya esté suspenso.
- En la última etapa de evaluación (escoger las recomendaciones), no se muestran los errores cometidos o el resultado obtenido.
- La nota final está dada de manera general (aprobado o no) sin hacer énfasis en los errores cometidos.
- No presenta una interfaz de usuario intuitiva.
- El usuario debe realizar numerosos toques en pantalla para completar una acción y, debido a que el sistema se utilizará en personas de avanzada edad, dificulta el manejo del sistema.
- Los colores del sistema son muy oscuros, lo que provoca que el usuario niegue la experiencia.

1.3.3. Nuevas Acciones Sugeridas

A partir de los señalamientos realizados anteriormente, se proponen un grupo de cambios o acciones a implementar en la nueva versión que se desea hacer de este sistema:

- Cada entrenamiento será dividido en tres etapas diferentes, que se evalúen en pruebas apartes (variables, causas y recomendaciones). Solo si se aprueban los entrenamientos de la primera etapa, se podrán realizar los de la segunda, y así sucesivamente.
- Se vincularán varios tipos de preguntas, para variar la forma de enfocar los pensamientos de los operadores y así contribuir a su aprendizaje.
- Las notas se verán por etapas, teniendo en cuenta el rendimiento de las pruebas así como la cantidad de intentos. La nota final será la suma de las tres etapas.
- Las visuales tendrán un cambio radical, con botones y carteles de ayuda para indicar al usuario lo que debe hacer.
- Se incorporarán imágenes en las visuales, para crear un ambiente amistoso e intuitivo.

■ El color predominante será el naranja. Esto se debe a que ese color ayuda en el subconsciente al aprendizaje.

1.4. Conclusiones Parciales

A partir de la investigación realizada y los aspectos profundizados a lo largo del capítulo:

- Se hizo un estudio detallado de los sistemas de capacitación existentes, llegando a concluir que, por las condiciones que presentan, el sistema de capacitación que más se ajusta a la Industria Alimentaria Cubana es el digital.
- Se tuvieron en cuenta todos los detalles de la problemática planteada y, gracias a las similitudes existentes, se concluyó que un sistema experto resulta la herramienta idónea para esta solución.
- Se analizó con profundidad el sistema SECPROIT para ser tomado como base a la hora de diseñar y programar el nuevo sistema. Será el modelo a seguir y se rectificarán todos las limitaciones existentes en él para esta nueva versión.

Capítulo 2

Nuevo Sistema de Entrenamiento

En este capítulo se presentará la propuesta de solución para este nuevo sistema de entrenamiento que se desea conseguir. Para que el desarrollo de un proyecto concluya con éxitos, primero debe realizarse un diseño completo de lo que se pretende obtener, así como los recursos y requisitos que debe poseer. A lo largo del capítulo se tratarán de presentar numerosos diagramas para facilitar la comprensión. Cabe destacar que este nuevo diseño parte del SECPROIT, por lo que se realizarán algunas comparaciones con ese sistema.

2.1. Modelado del Negocio

Los modelos se crean con el objetivo de entender mejor la entidad real que se va a construir. Deben cumplir con objetivos diferentes, niveles de abstracción, con la ilustración del software desde el punto de vista del cliente y después con su representación en un nivel más técnico.

2.1.1. Modelo de Dominio

El modelo de dominio es una representación de las clases conceptuales del mundo real, no de componentes de software. Se muestran los conceptos significativos en un dominio del problema. Para una mejor comprensión del modelo, es necesario entender las entidades que influyen en el negocio.

En la Figura 2.1, se puede apreciar el modelo de dominio para este nuevo sistema. Es muy semejante, en cuanto a lógica del negocio, al modelo del SECPROT.

Las áreas de color verde, indica que se ha modificado la entidad para ajustarla a los nuevos cambios que se decidieron hacer. Los cambios más significativos son aquellas entidades que se modificaron casi en su totalidad o aquellas que fueron agregadas de cero. Estas últimas están representadas con un color amarillo.



Figura 2.1: Modelo de Dominio del Nuevo Sistema.

Como bien se aprecia, el diagrama indica que se han creado cuatro nuevas entidades: configuración de entrenamiento, entrenamiento de variables, entrenamiento de causas y entrenamiento de recomendaciones. Estas entidades fueron creadas para un mejor manejo de la información de la capacitación, y para introducir una nueva regla: separar los entrenamientos por etapas.

2.1.2. Reglas del Negocio

Cada negocio posee un conjunto de restricciones que permiten su correcto funcionamiento. Estas restricciones son conocidas como reglas del negocio. Las reglas a cumplir en este nuevo sistemas son:

- Antes de crear un usuario, deben haberse creado los roles del sistema y las áreas laborales.
- Antes de crear un proceso, deben haberse creado las áreas laborales.
- Antes de crear un entrenamiento debe existir un proceso y al menos un operario.
- Antes de realizar el entrenamiento de las causas se debe haber aprobado el entrenamiento de las variables.
- Antes de realizar el entrenamiento de las recomendaciones se debe haber aprobado el entrenamiento de las causas.
- Los usuarios solo pueden poseer un rol.

- Los usuarios solo pueden pertenecer a un área a la vez.
- Los procesos solo pueden pertenecer a un área a la vez.
- Las tareas de los roles son únicas para el usuario, es decir, solo las pueden hacer los usuarios con el mismo rol.

2.1.3. Diagrama de Actividades

Un diagrama de actividades recoge una vista dinámica de alto nivel, del flujo principal del proceso. En el mismo, se tienen en cuenta las entidades que participan en el negocio y las acciones que realizan cada una, así como los llamados objetos con los que se trabaja. En estos diagramas se representa el inicio y el fin del flujo.

En esta propuesta de solución, el flujo del administrador resulta el mismo que el del sistema SECPROIT [8]. Sin embargo, el flujo de los operarios y jefes de área cambia. En la Figura 2.2, se representa el diagrama de actividades de estos dos roles, detallando el nuevo flujo de los mismos.



Figura 2.2: Diagrama de Actividades del Nuevo Sistema.

2.2. Captura de Requisitos

En la captura de requisitos se obtiene el modelo del sistema a desarrollar, quedando definidas las funcionalidades requeridas para establecer las fronteras del sistema. Además, se proponen otros elementos para lograr un mayor grado de especificación en el propio modelado del sistema.

2.2.1. Requisitos Funcionales

Los requisitos funcionales denotan las funcionalidades del sistema. Estos son capacidades o condiciones que se deben cumplir. Las funcionalidades del nuevo sistema son:

- Gestionar Usuarios: es una actividad desarrollada por los administradores. Se basa en introducir, modificar y desactivar usuarios en el sistema. También cuentan con una opción para restablecer la contraseña de los usuarios, en caso de que estos las hayan olvidado.
- Gestionar Áreas: es una actividad desarrollada por los administradores. Se basa en introducir, modificar y eliminar las áreas del sistema. Para poder eliminar un área laboral, esta debe estar vacía, es decir, que de ella no dependan ni procesos ni usuarios.
- Ver Acciones de los Usuarios: es una actividad desarrollada por los administradores, que les permite ver todas las actividades relacionadas con el sistema, realizadas por los usuarios. A modo de resumen, registra los camios realizados junto al nombre del usuario que lo hizo.
- Cambiar Contraseña: es una acción que todos los usuarios pueden realizar.
- Entrar en el Sistema: es una acción que todos los usuarios pueden realizar. Consiste en introducir las credenciales (nombre de usuario y contraseña) en el sistema.
- Gestionar Procesos: es una actividad desarrollada por los jefes de área. Se basa en introducir, modificar y eliminar los procesos en un área laboral. Si se elimina un proceso, queda registro del mismo y los entrenamientos que se hayan realizado no se pierden.
- Configurar Entrenamiento: es una actividad desarrollada por los jefes de área. Consiste en asignar para cada proceso, la cantidad de pruebas por etapas, los estilos de pregunta que se pueden aplicar, la cantidad general de intentos que se pueden realizar y los usuarios que pueden entrenar este proceso.
- Realizar Entrenamiento: es una acción desarrollada por los operarios. Consiste en responder las preguntas de los entrenamientos.

2.2.2. Requisitos No Funcionales

Los requisitos no funcionales son propiedades, características y restricciones que se deben presentar el sistema, asegurando el cumplimiento de las acciones demandadas. Los requisitos no funcionales de esta propuesta de solución son:

- Usabilidad: se debe garantizar un ambiente de trabajo simple e intuitivo, ya que la mayoría de los usuarios no poseen experiencias con sistemas informáticos.
- Seguridad: la información del sistema solo puede ser manipulada por usuarios autorizados (aquellos que posean usuario y contraseña). Las actividades deben estar divididas por roles, lo que presenta menor riesgo de errores.
- Confiabilidad: se debe evitar los enlaces rotos. La información contenida en los ficheros de los procesos, se debe validar antes de ser usada, para evitar errores.
 Se debe garantizar la confidencialidad de la información.
- Disponibilidad: la aplicación de mecanismos de seguridad no debe constituir un retraso para el uso del sistema. El software siempre debe estar disponible, así como brindar su información actualizada.
- Software: se debe tener instalado el JDK versión 1.8 y la aplicación PostgreSQL versión 9.1 (como mínimo). Esta última se utiliza como gestor de base de datos.
- Hardware: se necesitan 64MB de memoria RAM, un microprocesador Pentium II a 450 MHz como mínimo, y un disco duro con capacidad libre mínima de 4 GB, para almacenar la nueva información incorporada. Además, es necesario un sistema operativo de entorno gráfico como Windows y Linux y que estén establecidas las condiciones necesarias para el acceso al servidor.
- Portabilidad: debe ser utilizado bajo sistemas operativos Windows o Linux, por lo que su desarrollo debe realizarse con un lenguaje y tecnologías capaces de brindar este soporte. Además, el hecho de ser multiplataforma permite que el sistema llegue a un mayor número de personas.
- Restricciones en el diseño y la implementación: debe desarrollarse sobre plataformas de software libre y código abierto. Además, se utiliza la herramienta Drools, por lo tanto, el lenguaje de programación debe ser Java.
- Políticos/Culturales: debe encontrarse en idioma español.

2.3. Casos de Uso del Nuevo Sistema

Los casos de uso del sistema establecen un acuerdo entre clientes y desarrolladores sobre las condiciones y requisitos que debe cumplir la aplicación. Es un artefacto narrativo que describe, bajo la forma de acciones y reacciones, el comportamiento del sistema desde el punto de vista del usuario.

2.3.1. Actores del Sistema

Un actor del sistema es una entidad externa relacionada con el mismo y que demanda una funcionalidad. Esto incluye a los operadores humanos, pero también incluye a todos los sistemas externos. En la Tabla 2.1, se muestran todos los actores del nuevo sistema y una breve explicación de los mismos:

Actor	Descripción
Usuario	Es el actor genérico que hace uso de las fun-
	cionalidades que son comunes.
Operario	Es el actor que se evalúa a partir de los en-
	trenamientos de los procesos.
Jefe de Área	Es el actor que vincula los procesos en las
	áreas y los configura, decidiendo qué usua-
	rios se entrenarán y las características que
	deberán cumplir para aprobar. Anterior-
	mente era conocido como Especialista.
Administrador	Es el actor que gestiona las áreas, los usua-
	rios y visualiza las operaciones realizadas
	por los mismos.
Generador de Bases de Conocimientos	Es un sistema externo que genera las bases
	de conocimientos con las que se trabaja en
	los entrenamientos (procesos).

Tabla 2.1: Actores del Nuevo Sistema.

2.3.2. Diagrama de Casos de Uso del Nuevo Sistema

Un caso de uso es una secuencia de interacciones entre un sistema y alguien o algo que usa alguno de sus servicios. Describen tanto lo que hace el actor como la repuesta que recibe, aunque el énfasis está puesto en la interacción. Es iniciado por un actor y, a partir de ese momento, se intercambian datos. En la Figura 2.3, se puede observar el diagrama de casos de uso del nuevo sistema.

El diagrama, en general, tiene como base el sistema SECPROIT, pero con algunas modificaciones incluidas. Los casos de uso que se encuentran de color verde, indican que, en principio, realizan las mismas operaciones, pero con cambios incorporados. El actor del mismo color se debe a que cambió de nombre, de especialista a jefe de área. El caso de uso de color amarillo, es completamente nuevo.

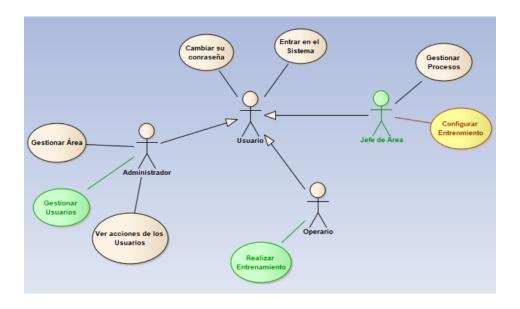


Figura 2.3: Diagrama de Casos de Uso del Nuevo Sistema.

2.4. Modelo de Datos

En el desarrollo de esta herramienta se utiliza: una bases de datos en PostgreSQL y un grupo de ficheros anm y drl que conforman los proceso de los entrenamientos. En la base de datos se almacena toda la información referente al sistema: los datos de los usuarios, los resultados de los entrenamientos, los procesos, la información de las áreas y una lista detallada de todas las operaciones realizadas en el sistema.

2.4.1. Diagrama de Entidades y Relaciones (DER)

Un Diagrama de Entidades y Relaciones (DER) representa las clases del sistema, con sus atributos, y las relaciones existentes entre ellas.

Referencias bibliográficas

- [1] Ángel Alegre Burcio, Implementación de estrategias de Computation Offloading en redes de sensores inalámbricas de bajo consumo para Internet de las Cosas. resreport, E.T.S.I. Industriales (UPM), Biblioteca ETSI Industriales, Feb. 2022.
- [2] I. D. Mateo and A. P. Salgueiro, "Vinculación de la formación doctoral con los programas nacionales de investigación en cuba. la experiencia del doctorado en cibernética aplicada del instituto de cibernética, matemática y física (icimaf)," Revista Control, Cibernética y Automatización, vol. 9, p. 61, Jan. 2021.
- [3] A. S. García, "Trabajo en la agroindustria cañera cubana, 1898-1914: estrategias para aumentar la producción con escasez de mano de obra," Consejo Superior de Investigaciones Científicas (España), pp. 171–206, Apr. 2022.
- [4] F. S. Concepción, R. P. Aguilar, A. P. Moreno, A. A. Cordoví, J. C. Castro, and O. L. Santiago, "Metodología para la automatización de procesos tecnológicos en la industria farmacéutica cubana," *Ingeniería Industrial*, vol. 43, Feb. 2022.
- [5] T. de las Mercedes Escoriza Martínez, "Aplicación de un procedimiento para la gestión de riesgos en procesos de la industria cárnica cubana," I CONFERENCIA INTERNACIONAL DE INGENIERÍA INDUSTRIAL (CINDUS 2019), vol. II, p. 13, June 2019.
- [6] G. C. Lemus, R. S. Llanes, and A. M. R. Riverón, "Sistema para gestionar bases de conocimiento en la industria alimentaria," Revista Cubana de Ingeniería, vol. IX, pp. 60 – 68, May 2018.
- [7] G. C. L. Ana Mailen Rodríguez Riverón, Raisa Socorro Llanes, SISTEMA PARA GESTIONAR BASES DE CONOCIMIENTO EN LA INDUSTRIA ALIMENTA-RIA. software, Universidad Tecnológica de la Habana "José Antonio Echeverría", Oct. 2017.
- [8] R. S. L. Elena Acosta Gil, SISTEMA PARA EL ENTRENAMIENTO DE OPERA-RIOS EN PROCESOS TECNOLÓGICOS EN LA INDUSTRIA ALIMENTARIA.

- software, Universidad Tecnológica de la Habana "José Antonio Echeverría", June 2018.
- [9] "Cómo crear un buen plan de capacitación laboral," *Bizneo Blog*, vol. 1, p. 5, June 2022.
- [10] I. de Seguridad Minera ISEM, "Entrenamiento basado en computador para operadores y mantenedores de planta," Seguridad Minera, June 2022.
- [11] E. Salado, "Los 6 mejores software de capacitación de personal y formación para empleados," *SoftDoit*, June 2022.
- [12] K. H. C. Rogelio E. Martínez R., Guillermo Romero J., Sistema de capacitación asistido por computadora, accesible vía Internet. Aplicaciones tecnológicas, June 2002.
- [13] G. Aretio, "Etapas de la evaluación de calidad de los aprendizajes," Contextos universitarios mediados, p. 3, May 2020. García es un Catedrático Emérito de la UNED. Director/Editor de la RIED. Revista Iberoamericana de Educación a Distancia. Presidente de Honor Cátedra UNESCO de Educación a Distancia (CUED).
- [14] E. V. Laguna, "Metodología estándar para el entrenamiento básico de un panel de catadores," candthesis, Universidad: Rey Juan Carlos, July 2016.
- [15] G. W. Philipp Christmann, Rishiraj Saha Roy, "Beyond ned: Fast and effective search space reduction for complex question answering over knowledge bases," Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining, pp. 172–180, Feb. 2022.
- [16] C. M. J. Alty J L, Expert systems: concepts and examples. John Wiley and Sons Inc., New York, NY, Jan. 1984. 99 GENERAL AND MISCELLANEO-US//MATHEMATICS, COMPUTING, AND INFORMATION SCIENCE; EXPERT SYSTEMS; REVIEWS; ARTIFICIAL INTELLIGENCE; C CODES; COMPUTER CALCULATIONS; COMPUTER CODES; DATA PROCESSING; I CODES; M CODES; P CODES; R CODES; DOCUMENT TYPES; PROCESSING; 990200.
- [17] M. J. B. Ronald L. Jacobs, "Developing employee expertise through structured onthejob training (sojt): an introduction to this training approach and the knpc experience," *Industrial and Commercial Training*, vol. 44, pp. 75–84, Mar. 2012.

- [18] F. Rasheed and A. Wahid, "Learning style detection in e-learning systems using machine learning techniques," Expert Systems with Applications, vol. 174, p. 114774, 2021.
- [19] S. J. S. Kandula Neha, "Analysis of student academic performance through expert systems," *International Research Journal on Advanced Science Hub*, vol. 2, no. Special Issue ICIES 9S, pp. 48–54, 2020.
- [20] A. V. Potapova, G. S. Tibilova, A. V. Ovcharenko, and N. V. Diachenko, "Designing a network of expert systems for identifying recipients of public services," in Knowledge in the Information Society (D. Bylieva, A. Nordmann, O. Shipunova, and V. Volkova, eds.), (Cham), pp. 136–148, Springer International Publishing, 2021.
- [21] E. Odhiambo Omuya, G. Onyango Okeyo, and M. Waema Kimwele, "Feature selection for classification using principal component analysis and information gain," *Expert Systems with Applications*, vol. 174, p. 114765, 2021.