



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERIA INDUSTRIAL

**DISEÑO Y CONSTRUCCIÓN DE UN SISTEMA DE SUGERENCIAS DE
PREGUNTAS PAES DE MATEMÁTICAS BASADO EN ALGORITMOS
DE RECOMENDACIÓN DE CONTENIDO RELEVANTE**

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN GESTIÓN DE OPERACIONES

ROMINNA ANDREA JIMÉNEZ OLIVERA

**PROFESOR GUÍA:
JUAN VELÁSQUEZ SILVA**

**MIEMBROS DE LA COMISIÓN:
VÍCTOR HERNÁNDEZ MARTÍNEZ
CHARLES THRAVES CORTÉS-MONROY**

**SANTIAGO DE CHILE
2024**

DISEÑO Y CONSTRUCCIÓN DE UN SISTEMA DE SUGERENCIAS DE PREGUNTAS PAES DE MATEMÁTICAS BASADO EN ALGORITMOS DE RECOMENDACIÓN DE CONTENIDO RELEVANTE

El objetivo de este trabajo de tesis, con el fin de mejorar el acceso a todo tipo de estudiantes a la ejercitación de la Prueba de Acceso a la Educación Superior de matemáticas, es "Diseñar, construir e implementar un sistema de recomendación de preguntas PAES de matemáticas que se base en lo que el usuario va respondiendo (que aprenda de él) para realizar recomendaciones de las siguientes preguntas que se le realizarán, de manera de atacar sus puntos más débiles y así poder ejercitar lo que más necesita mejorar".

Actualmente por falta de conocimiento o de tiempo los apoderados no pueden entregar una dedicación completa a sus pupilos y, por otro lado, los estudiantes ejercitan sin tener una guía adecuada o no ejercitan lo suficiente.

Una solución que ataca directamente este problema son las clases particulares, que tienen un costo asociado, generalmente elevado. Otras alternativas para resolverlo son plataformas E-Learning, o laboratorios, por ejemplo, Squirrel AI. En el caso de las plataformas E-Learning estas incorporan funcionalidades mediante las cuales se tiene acceso a material síncrono y asíncrono que puede ayudar a reforzar los contenidos, pero sin una guía que pueda dirigir el aprendizaje. Por otro lado, en el caso del laboratorio Squirrel AI, se cumple bastante a cabalidad el objetivo que se busca, dar la dedicación que cada alumno necesita y de manera personalizada, pero requiere de la presencia de tutores que están de manera online o presencial en los laboratorios. A raíz de esto nace como hipótesis de investigación que "los estudiantes pueden obtener mejores resultados en su PAES de matemáticas si utilizan como ejercitación el método de Kuizus"

Específicamente se propone un sistema, llamado Kuizus, el cual incorpora algoritmos de recomendación que buscan optimizar la cantidad y orden en que se contestan las preguntas tipo PAES de matemáticas, personalizando ambas variables para las necesidades de cada usuario.

Para probar la hipótesis de investigación se llevó a cabo una metodología que incorpora técnicas de recomendación basándose en las respuestas previas de los usuarios, para realizar las recomendaciones, pudiendo comparar el diagnóstico inicial con el diagnóstico final. Para poder comprobar si la hipótesis de investigación (que haya mejora en los alumnos utilizando el sistema de recomendación) se cumple, se consideró a 22 estudiantes de tercero y cuarto medio, donde 11 fueron parte del grupo de control (luego del diagnóstico se les realizaron preguntas aleatorias como entrenamiento), y 11 fueron parte del grupo experimental (su entrenamiento era con preguntas recomendadas por el sistema).

Los resultados del experimento mostraron que el grupo experimental tuvo una mejora de un 52% mayor que el grupo de control. El grupo experimental tuvo una mejora de 14,8% versus la mejora de 9,7% del grupo de control, por lo que se comprobó que la hipótesis de investigación se cumple. Aunque, como se puede ver, ambos grupos mejoraron su desempeño, por lo que se puede concluir que con el sólo hecho de aumentar la ejercitación se mejoran los resultados, de todas maneras puede observarse que sí hay un factor que aumenta en mayor parte la mejora, al implementar el algoritmo, versus simplemente ejercitar de manera aleatoria, por lo que se vuelve muy relevante como trabajo futuro poder mejorar aún más el algoritmo incorporando la técnica que involucra recomendaciones basadas en los pares similares, que esta vez no pudo ser incorporada por el bajo número de personas que utilizaron el sistema.

Tabla de Contenido

Capítulo 1: Introducción	1
1.1 Descripción del Proyecto.....	1
1.1.1 Fenómeno a observar	1
1.1.2 Dolor.....	2
1.1.3 Problema.....	2
1.1.4 Cliente.....	2
1.1.5 Usuario	2
1.2 Contexto e Hipótesis de Investigación	2
1.3 Objetivo General	3
1.4 Objetivos Específicos.....	3
1.5 Metodología.....	4
1.6 Alcances	4
1.7 Contribuciones	4
Capítulo 2: Marco Teórico	5
2.1 Sistemas de Recomendación	5
2.1.1 Dimensiones de los Sistemas de Recomendación	5
2.1.2 Técnicas de Recomendación	8
2.1.3 Indicadores en Recomendaciones.....	13
2.1.4 Funciones de Similitud.....	17
2.1.5 Medidas de Similitud	18
2.1.6 Indicadores de Evaluación de Sistemas de Recomendación	23
2.2 Plataformas E-Learning.....	24
2.2.1 Características Plataformas E-Learning	25
2.2.2 Sistemas de Recomendación para E-Learning.....	29
2.3 Caso de éxito: Squirrel AI	30
2.3.1 Análisis de Metodología de Squirrel AI.....	31
2.4 Análisis Funcional de Plataformas Similares.....	32
Capítulo 3: Levantamiento de Requerimientos.....	34
3.1 Levantamiento de <i>Features</i> Mínimas.....	34
3.1.1 Sistemas de Recomendación	34

3.1.2	Plataformas E-learning	36
3.1.3	Squirrel AI	38
3.2	Situación Actual	38
3.3	Diseño Conceptual	39
Capítulo 4:	Propuesta de Solución.....	40
4.1	Propuesta de Solución	40
4.2	Modelo de Proceso	41
4.2.1	Usuario Ingresa a la Plataforma	41
4.2.2	Usuario Inicia Sesión.....	41
4.2.3	Usuario Se Registra	41
4.2.4	Usuario Verifica su Cuenta.....	42
4.2.5	Inicio de la Plataforma.....	42
4.2.6	Comenzar a Ejercitar con Diagnóstico.....	42
4.2.7	Ejercitación.....	42
4.2.8	Usuario Está Ejercitando.....	42
4.2.9	Pantalla de Inicio	43
4.3	Categorización y Obtención de Preguntas PAES.....	44
4.4	Base para las Recomendaciones.....	45
4.4.1	Input Implícito	45
4.4.2	Asociación de Items.....	46
4.4.3	Similitud entre Usuarios.....	46
4.5	Algoritmo Recomendador	47
4.5.1	Pseudo Código del Algoritmo	48
4.5.2	Ejemplo de Funcionamiento del Algoritmo	48
4.6	Arquitectura a Implementar	52
4.6.1	Aplicación Web.....	52
4.6.2	Motor Recomendador	52
4.6.3	Base de Datos	53
4.7	Modelo de Datos	54
4.8	Modelo de Interfaz	55
4.8.1	Vista de Diagnóstico.....	55
4.8.2	Vista de Entrenamiento.....	56

4.9 Tecnologías Utilizadas.....	57
Capítulo 5: Evaluación de Sistema Kuizus.....	60
5.1 Diseño de Evaluación	60
5.2 Validación de la Evaluación.....	61
5.3 Metodología de la Evaluación.....	62
5.3.1 Grupo Experimental y Grupo de Control	62
5.3.2 Parámetros a Considerar	62
Capítulo 6: Resultados	64
6.1 Resultados Obtenidos	64
6.1.1 Diagnóstico Inicial.....	64
6.1.2 Entrenamiento	65
6.1.3 Diagnóstico Final.....	66
6.2 Análisis de Resultados	67
Capítulo 7: Conclusiones	71
7.1 Acerca de la Hipótesis de Investigación	71
7.2 Acerca de la Metodología Utilizada en el Algoritmo Recomendador	71
7.3 Acerca de los Objetivos.....	72
7.4 Trabajo Futuro	72
Capítulo 8: Bibliografía.....	74
Anexos	77
Anexo A: Temario PAES Matemáticas	77
Anexo B: Algoritmo de Sistema Recomendador.....	80
Anexo C: Extracto Distribución t de Student.....	80

Índice de Tablas

Tabla 1: Resumen de Enfoques de Técnicas de Recomendación - Elaboración Propia	10
Tabla 2: Ejemplo de Conjuntos - Elaboración Propia	19
Tabla 3: Análisis de Funcionalidades - Elaboración Propia	32
Tabla 4: 1° iteración de tabla de puntajes de los usuarios – Elaboración Propia	50
Tabla 5: 2° iteración de tabla de puntajes de los usuarios – Elaboración Propia	51
Tabla 6: Resultado Preguntas Diagnóstico Inicial - Elaboración Propia	64
Tabla 7: Preguntas realizadas por tema en Entrenamiento - Elaboración Propia	65
Tabla 8: Resultado Preguntas Entrenamiento - Elaboración Propia	66

Tabla 9: Resultado Preguntas Diagnóstico Final - Elaboración Propia.....	66
--	----

Índice de Figuras

Figura 1: Explicabilidad Versus Accuracy, - Basado en “Accuracy-Model Interpretation Trade-Off” [11] ..	15
Figura 2: Esquema de ejemplo de Similitud de Cosine	21
Figura 3: Proceso del Usuario en la Plataforma - Elaboración Propia.....	43
Figura 4: Proceso de Interacción Aplicación Web y Motor Recomendador - Elaboración Propia.....	53
Figura 5: Diagrama Entidad Relación - Elaboración Propia.....	54
Figura 6: Ejemplo de Vista de pregunta del diagnóstico - Elaboración Propia	56
Figura 7: Ejemplo de Vista de pregunta del entrenamiento - Elaboración Propia	57
Figura 8: Arquitectura Tecnológica - Elaboración Propia.....	59
Figura 9: Número de Preguntas por Tema - Elaboración Propia	67
Figura 10: Respuestas correctas de los diagnósticos del grupo de control - Elaboración Propia.....	68
Figura 11: Respuestas correctas de los diagnósticos del grupo experimental - Elaboración Propia.....	69

Capítulo 1: Introducción

Este trabajo se desarrolla con el fin de poder validar el funcionamiento y desempeño de un sistema de recomendación de preguntas PAES de matemáticas.

1.1 Descripción del Proyecto

Con el fin de mejorar el acceso a todo tipo de estudiantes a la ejercitación de la Prueba de Acceso a la Educación Superior (PAES) de matemáticas, se busca desarrollar un sistema de aprendizaje, Kuizus, que, mediante distintos tipos de ejercicios logre capturar qué es lo que el estudiante no sabe, de manera que, para reforzar el contenido que está más débil en su entendimiento, el sistema lo ayude sugiriéndole preguntas con las temáticas que él no comprende, en base a lo que se está detectando como los ejes temáticos o habilidades en los cuales ocurren la mayor cantidad de errores.

De esta forma, el sistema le recomendará problemas que serán más desafiantes en base a su comportamiento anterior, es decir, preguntas en los temas que demuestra tener deficiencia, en base a las respuestas que dio a los ejercicios presentados por el sistema anteriormente, consiguiendo así, progreso en el ámbito en que está con más dificultades.

1.1.1 Fenómeno a observar

Actualmente dada la situación del país, donde el promedio de los cursos tiene 31 estudiantes [1] en una sala de clases y el 30% del total tiene más de 34¹ se dificulta bastante poder hacer un seguimiento personalizado al aprendizaje individual de cada niño. Además, no todos pueden acceder a

¹ de acuerdo a https://accioneducar.cl/wp-content/files_mf/1443578310Ana%CC%81lisisAlumnosporCurso_26.6.2015.pdf, el promedio ponderado de las medianas según tipo de establecimiento de la cantidad de alumnos por curso es 31, donde el 30% tiene más de 34 alumnos por curso.

clases particulares en caso de quedar atrás con las materias que se están viendo en el colegio.

1.1.2 Dolor

Por falta de conocimiento o bien por falta de tiempo los apoderados, actualmente no pueden dedicarse tanto como les gustaría a ayudarles a estudiar a sus hijos. Por otro lado, los estudiantes ejercitan sin tener una guía adecuada que los haga seguir el camino en el cuál deberían enfocarse.

1.1.3 Problema

Si los niños no estudian o no ejercitan lo suficiente obtendrán notas deficientes, lo que en el tiempo podría transformarse en un fracaso académico, impidiéndolos de estudiar lo que les gustaría y así trabajar y tener una buena situación en el futuro.

1.1.4 Cliente

- Colegios, pensando en que el rendimiento de sus estudiantes mejore.
- Apoderados, considerando el desempeño y bienestar de sus hijos.
- Postulantes a educación superior, quienes buscan obtener un buen resultado en la PAES.

1.1.5 Usuario

Los estudiantes, serán quienes utilicen el sistema para poder tener una guía adecuada que les permita ejercitar de mejor manera.

1.2 Contexto e Hipótesis de Investigación

Considerando lo levantado en el punto "Fenómeno a Observar" no queda mucho espacio para la dedicación exclusiva al aprendizaje de cada

estudiante, y como se mencionó también, acceder a clases particulares puede resultar costoso (entre \$15.000 y \$25.000 por hora²).

Hipótesis: “Los estudiantes pueden obtener mejores resultados en su PAES de matemáticas si utilizan como ejercitación el método Kuizus.”

Específicamente se propone un sistema que incorpora algoritmos de recomendación que buscan optimizar la cantidad y orden en que se contestan las preguntas, personalizando ambas variables para las necesidades de cada usuario.

1.3 Objetivo General

Diseñar, construir e implementar un sistema de recomendación de preguntas PAES de matemáticas que se base en lo que el usuario va respondiendo (que aprenda de él) para realizar recomendaciones de las siguientes preguntas que se le realizarán, de manera de atacar sus puntos más débiles y así poder ejercitar lo que más necesita mejorar.

1.4 Objetivos Específicos

- Revisar estado del arte actual respecto a plataformas e-learning y sistemas de recomendación.
- Definir la estructura del sistema de recomendación a utilizar.
- Definir las técnicas de recomendación a utilizar.
- Construir el sistema recomendador incorporando las técnicas de recomendación definidas.
- Validar funcionamiento del sistema recomendador.

² www.tusclasesparticulares.cl

1.5 Metodología

- Estudiar la literatura existente en algoritmos de recomendación y plataformas e-learning, esto considera también distintos sistemas de recomendación.
- Entender la forma en que las preguntas PAES son categorizadas respecto a los distintos temas que abarcan y comprender como se relacionan entre sí.
- Aplicar el conocimiento de la literatura en sistemas de recomendación y plataformas e-learning para crear algoritmo de recomendación
- Incorporar las técnicas de recomendación seleccionadas al sistema recomendador.
- Realizar evaluación para validar el funcionamiento del sistema recomendador.

1.6 Alcances

El proyecto que se presenta considera la creación de un sistema de ejercitación en una plataforma web que incluye algoritmos de recomendación para la ejercitación de preguntas de matemáticas, específicamente tipo PAES.

1.7 Contribuciones

Obtención de un sistema recomendador que hará preguntas de matemáticas tipo PAES al usuario. El sistema tendrá memoria y aprenderá del usuario, ya que, dependiendo de las respuestas de éste, será la siguiente pregunta que le realizará. El sistema aprenderá con el tipo de respuestas que se entregan, ya que sabrá en que se está equivocando el usuario dado lo que respondió. Con esta información podrá realizar preguntas que se centren en las temáticas en las que el usuario está cometiendo errores o no está entendiendo bien, para así poder dedicarle mayor ejercitación y mejorar en estos ámbitos más débiles.

Dado esto, además de esta tesis (la documentación del sistema), el entregable es el sistema recomendador de preguntas PAES de matemáticas habilitado en una plataforma web para poder ser utilizado.

Capítulo 2: Marco Teórico

Lo que más destacaba en la búsqueda de sistemas de recomendación para el aprendizaje en el ámbito educacional en relación a las recomendaciones eran algoritmos de recomendación, pero utilizados en otros ámbitos, por ejemplo, los algoritmos utilizados para la recomendación de películas o series en Netflix³ [2]. En el caso del aprendizaje en el ámbito educacional había bastantes plataformas *e-learning*⁴ de donde poder levantar los requerimientos que debería tener un sistema de recomendación con finalidad de aprendizaje. Además, un caso de un sistema bastante parecido a lo que Kuizus busca hacer, un sistema de inteligencia artificial chino, Squirrel AI.

En este capítulo se profundizará más respecto a estos tres temas para poder comprender de mejor manera la metodología utilizada en este trabajo.

2.1 Sistemas de Recomendación

En la literatura hay distintos tipos de sistemas de recomendación, que a su vez utilizan distintos tipos de técnicas, indicadores, y que también tienen distintas dimensiones para poder ser definidos [3].

2.1.1 Dimensiones de los Sistemas de Recomendación

Ellas son dominio, contexto, propósito, nivel de personalización, opiniones expertas, confianza y privacidad, interface, y algoritmos. A continuación, la descripción de cada una de ellas:

- Dominio: es el tipo de contenido recomendado, en el ejemplo de Netflix sería películas y series [3], en el caso de este trabajo el dominio serían las preguntas de matemáticas que se están recomendando.

³ Sitio de *streaming*, principalmente de películas y series.

⁴ Espacio virtual de aprendizaje.

- Contexto: es el ambiente en el cual el usuario recibe la recomendación, en el caso ejemplo de Netflix sería el dispositivo mediante el cual ven el *streaming* [3] en este caso sería el *browser*⁵ que se esté utilizando para acceder a la plataforma.
- Propósito: es el fin que tiene el sistema recomendador, en el caso de Netflix, hay un propósito tanto para el proveedor como para el usuario final. En el primer caso el propósito sería que los clientes paguen por la suscripción mediante el contenido que Netflix le brinda, y para el usuario final su propósito es encontrar contenido relevante que desee ver en un momento específico [3] . En el caso del sistema recomendador de preguntas PAES el propósito es que el usuario pueda mejorar su entendimiento y destreza respecto a las temáticas que incorporan.
- Nivel de personalización: las recomendaciones pueden tener distintos niveles de personalización. Por ejemplo, no personalizado viene de usar estadísticas generales, al final todos los usuarios reciben las mismas recomendaciones. El semi personalizado: personaliza por segmento (país, edad, gente que maneja, gente que camina), y el personalizado se basa en información con respecto al usuario que indica cómo ha interactuado anteriormente con él [3]. En este caso se busca nivel personalizado ya que las recomendaciones se realizarían basadas en el comportamiento del usuario.
- Opiniones expertas: no se utilizan tan frecuentemente como en el pasado, pero aún hay algunas que prevalecen por ejemplo en el caso de los vinos, que como el nombre de la dimensión lo señala, pueden ser recomendados por un experto [3]. En este caso no se utilizarían, sino que, al momento de la recomendación, estas se basarían en las respuestas entregadas por el usuario a las preguntas anteriores, de manera de poder ir construyendo un perfil y hacer las recomendaciones de manera personalizada.

⁵ Aplicación de *software* que sirve para localizar, recibir y mostrar con tenido de internet.

- **Confianza y privacidad de las recomendaciones:** se basa en que los usuarios no se sientan manipulados y que confíen en las recomendaciones [3]. Puede que esta dimensión se dé más en los casos que involucren ventas, ya que los usuarios podrían sentir que la parte que quiere venderles podría recomendarles cualquier cosa, no necesariamente algo que sea conveniente para ellos, ya que sólo quieren venderles algo. En este caso es más difícil que los usuarios desconfíen del sistema, ya que en la plataforma siempre estarán recibiendo preguntas para ejercitarse, la diferencia sólo sería que después de conocer las respuestas que se están dando, se podría caracterizar su comportamiento para generar recomendaciones exclusivas de preguntas para el usuario, por lo que sólo podría ser beneficioso para el usuario. En estricto rigor al proveedor del sistema no le favorece ni le perjudica entregar o no recomendaciones, ya que el usuario ya estaría utilizando la plataforma que es el fin del proveedor, y el hecho de entregar recomendaciones es una ventaja para el usuario, ya que no sería un sistema de ejercitación tradicional, sino que se adecuaría a la realidad de conocimiento de cada usuario.
- **Interface:** describe los tipos de inputs y outputs que se producen. Un tipo de input explícito podría ser la propia opinión del usuario, que el mismo usuario agrega de forma manual. Un tipo de input implícito sería donde el sistema trata de deducir la opinión o el gusto mediante la forma de interacción con el mismo, pero sin que éste lo indique directamente. Los tipos de outputs pueden ser predicciones, recomendaciones o filtros. Si las recomendaciones son una parte natural de la página se dice que tienen presentación orgánica, si no, que son inorgánicas [3]. En este caso el input sería implícito ya que el usuario no diría explícitamente por ejemplo “se me dificultan las preguntas de isometría que evalúan habilidades de comprensión”, sino que mediante las respuestas que se dieran a las preguntas el sistema de recomendación inferiría cuáles son los ejes temáticos y habilidades que están más débiles en cada persona para poder reforzarlo con las recomendaciones. El output serían recomendaciones, más específicamente preguntas recomendadas.

El output como se acaba de mencionar serían las recomendaciones, que podrían tener una presentación orgánica o inorgánica [3]. Presentación orgánica quiere decir que son parte natural de la página, a diferencia del caso inorgánico. En el ejemplo de Netflix, este tiene una presentación orgánica, ya que no indica que las películas que se muestran son recomendaciones, sino que son una parte integral de la plataforma. Distinto sería por ejemplo un correo dirigido a alguna persona en particular, de algún cine que envíe la lista de las "mejores películas para ti", ya que en este caso las recomendaciones serían personalizadas e inorgánicas.

- Algoritmos: describe el tipo de algoritmos que se van a utilizar para realizar las recomendaciones [3]. Estos según la técnica que se utilice para recomendar pueden ser de distintos tipos. Se describen a continuación en la sección 2.1.2.

2.1.2 Técnicas de Recomendación

De acuerdo a la fuente de donde se obtiene la información para poder realizar las recomendaciones, las técnicas de recomendación pueden clasificarse en distintos tipos. Estos son colaborativas, demográficas, basadas en contenido, basadas en utilidad y basadas en conocimiento.

- Colaborativas: agrega los ratings de recomendaciones de objetos, reconoce lo que los usuarios que entregaron estas recomendaciones tienen en común, basándose en los ratings que dan a cada uno de los objetos y genera nuevas recomendaciones que se basan en comparaciones entre los usuarios.
- Demográficas: apunta a categorizar al usuario en base a sus atributos personales y hace recomendaciones en base a su clase demográfica.
- Basadas en contenido: los objetos de interés se definen por sus características asociadas. Un recomendador basado en contenido se aprende el perfil de los intereses de un usuario basado en las

características que presenta en los objetos a los que el usuario les ha dado una valoración. El tipo de perfil de usuario que se deriva de un recomendador basado en contenido depende del método de aprendizaje empleado (se usan arboles de decisión, redes neuronales y representaciones basadas en vectores).

Al igual que en el caso colaborativo los perfiles basados en contenido son modelos de largo plazo y se actualizan a medida que se observan mayor cantidad de preferencias.

- Basadas en utilidad: estas sugerencias se basan en un cómputo de la utilidad de cada objeto por usuario. El problema es cómo crear la función de utilidad para cada usuario. El perfil del usuario es la función de utilidad que el sistema ha derivado para el usuario empleando restricciones de satisfacción para localizar el mejor match. El beneficio de este tipo de recomendación es que puede ponderar características que no son propias del producto, por ejemplo, la confiabilidad del vendedor, o la disponibilidad del producto.
- Basadas en conocimiento: este tipo de recomendaciones apunta a sugerir objetos basados en inferencias respecto a necesidades y preferencias del usuario. En algún sentido todas las técnicas de recomendación pueden ser descritas como hacer algún tipo de inferencia. Estos tipos de recomendación se distinguen porque tienen conocimiento funcional, tienen conocimiento respecto a cómo un ítem en particular cumple con una necesidad específica de un usuario y puede razonar acerca de la relación entre una necesidad y una posible recomendación. El perfil del usuario puede ser cualquier estructura de conocimiento que apoye/soporte esta inferencia. En el caso más simple, como Google, puede ser simplemente la *query* que el usuario formula, en otros casos puede ser una representación más detallada de las necesidades del usuario. Este tipo de recomendaciones puede tener distintas formas, por ejemplo, Google usa información acerca de los links que hay entre páginas web para inferir popularidad y valor autoritativo. [4]

A diferencia de los tres primeros casos, tanto las recomendaciones basadas en utilidad, como las basadas en conocimiento no buscan construir generalizaciones a largo plazo con respecto a sus usuarios, sino que basar su recomendación en un match entre las necesidades del usuario y entre un conjunto de opciones disponibles. A continuación, una tabla resumen de cada enfoque:

Tipos de Técnicas de Recomendación	Colaborativa	Demográfica	Basadas en Contenido	Basadas en Utilidad	Basadas en Conocimiento
Categoriza al usuario basándose en su perfil		X			
Aprende perfil del usuario basándose en características de objetos que le gustan			X		
Reconoce usuarios en común basándose en los ratings entregados	X				
Puede ponderar características que no son intrínsecamente parte del producto				X	
Sugiere objetos en base a inferencias en las preferencias del usuario		X			X
Puede razonar acerca de la relación entre una necesidad y una posible recomendación					X

Tabla 1: Resumen de Enfoques de Técnicas de Recomendación - Elaboración Propia

Hay otras consideraciones que se pueden tener respecto a las técnicas de recomendación a utilizar, además de la fuente de donde provengan los datos, y estos son los elementos externos a tener en cuenta al momento de recomendar, uno de ellos son las reglas de asociación.

- Reglas de Asociación: muchas veces en vez de recomendar el objeto más popular, las recomendaciones se basan en la idea de las reglas de asociación. Por ejemplo, pensando en una persona que se está comprando un disco duro en alguna tienda, probablemente podría incluso haber llegado a su casa antes de darse cuenta que necesitaba un cable para poder conectarlo, pero en el caso de Amazon, como utiliza reglas de asociación este problema no ocurre [5], ya que cuando se está efectuando la compra, Amazon sugiere comprar un cable dado que se está comprando un disco duro. Esto ocurre, ya que en ocasiones anteriores usuarios que han comprado un disco duro también incluyen un cable en su compra, por lo que está considerando que, si una persona compra un disco duro, probablemente también va a necesitar un cable.

Para el caso de este sistema de recomendación no es exactamente como se presenta en los ejemplos, pero si con una variación, y es que probablemente hay contenidos y/o habilidades que están asociados entre ellos y se da la situación de que una persona que usualmente se equivoca en una pregunta de contenido a, también comete errores en preguntas de contenido b. Es algo que habría que estudiar a medida se vaya teniendo más información en el sistema, pero es importante considerarlo para agregarlo como un posible elemento para tener en cuenta al momento de hacer las recomendaciones.

Al igual que en el momento de revisar la posibilidad de incluir reglas de asociación si es que estas se dan en el contexto en que se desea realizar la recomendación hay que fijarse en otros elementos, como por ejemplo lo que se mencionaba en el punto anterior, "es algo que habría que estudiar a medida que se vaya teniendo más información en el sistema", el nuevo elemento sería poder contestar la pregunta, ¿pero cuándo va a ser suficiente información?, ¿Cuántos elementos tiene que rankear (preguntas tiene que resolver, en este caso) el usuario para poder comenzar a darle recomendaciones?

A un usuario se le llama *cold visitor* cuando no se sabe nada de él ni su comportamiento [6]. Entonces respondiendo la pregunta anterior, ¿Cuándo se le podrían empezar a dar recomendaciones personalizadas? Hay muchos paper científicos que dicen que las recomendaciones no se pueden calcular antes de que el usuario haya calificado al menos de 20 a 50 elementos, pero normalmente el usuario espera que se le entreguen recomendaciones mucho antes de eso [6]. Para esto se le pueden entregar recomendaciones no personalizadas desde el principio, y luego de tener la suficiente cantidad de elementos calificados por el usuario comenzar a entregarles recomendaciones personalizadas.

En el caso del sistema recomendador de esta memoria, se podría establecer un número base de preguntas aleatorias para así poder obtener el comportamiento inicial del usuario y que no sea un *cold visitor*, y desde ahí poder empezar a entregarle recomendaciones ya que ya se ha podido caracterizar su comportamiento con las respuestas a las preguntas iniciales que tuvo que resolver. Esto sería así (input implícito) porque a las recomendaciones en sistemas educacionales no le sirven las calificaciones explícitas [7], sería tal vez contraproducente preguntarle al usuario cuáles son las preguntas que no saben o que se les hacen más difíciles, muchas veces porque quizás realmente no tienen claridad al respecto, o porque podría ser una ventana ciega⁶.

Así mismo como en lo recién mencionado respecto al input explícito, sucede también el fenómeno contrario en los sistemas de recomendación educacionales en cuanto al concepto conocido como *time based approach*. En los sistemas de recomendación, generalmente los ítems más viejos son penalizados, esto ocurre porque de acuerdo a la experiencia en los sistemas de recomendación mientras más antigua sea una apreciación entregada por un usuario, más "puntos" pierde al momento de hacerle una nueva recomendación, por ejemplo una persona que siempre ha visto comedias románticas, y sólo una película de acción, pero da la casualidad de que fue la

⁶ En la Ventana de Johari (libro de psicología cognitiva), hace referencia a lo que una persona desconoce de si misma, pero que si es conocido por el resto.

última, ésta es la que debería considerarse mayormente al realizar las nuevas recomendaciones [8].

En el caso del sistema de recomendación de este trabajo ocurriría lo contrario de lo que se menciona en los sistemas de recomendación tradicionales como se da en el ejemplo de las películas del párrafo anterior. En este caso pueden darse dos situaciones, el primero puede ser que la pregunta se haya respondido de manera incorrecta, en este caso no es relevante, ya que como la respondió mal este es el tema que se vuelve "prioridad" en lo que a recomendaciones se refiere, sin embargo en el caso de que esta pregunta se haya respondido correctamente, a medida que se avance en el tiempo respecto a los temas de esta pregunta, por mucho que se haya respondido correctamente en su momento, se hace necesario según el *time-based approach*, considerar el tema de "lo más reciente es lo que vale". Para este caso como se menciona en el caso de que la respuesta sea incorrecta, los pasos a seguir son relativamente sencillos ya que dado que la pregunta se respondió mal, el eje temático y/o habilidad de la pregunta inmediatamente pasa a estar primero para la recomendación, pero en el caso en que se responda bien no implica que el usuario va a saber de esto para siempre, de hecho aplicando el *time-based approach*, como lo más reciente es lo que tiene mayor valor, son las preguntas que han respondido correctamente en los últimas instancias las que podrían omitirse al momento de la recomendación ya que se entenderían como "dominadas", sin embargo las preguntas que se respondieron correctamente hace más iteraciones deberían ir perdiendo aprendizaje o memoria, de manera que si ha pasado mucho tiempo en el que no se ha respondido una pregunta que tiene cierto eje temático y evalúa cierta habilidad, la recomendación debería ser entregar una pregunta de ese tipo al usuario para poder comprobar si aún sabe o no.

2.1.3 Indicadores en Recomendaciones

Hay indicadores que son muy relevantes al momento de comenzar a hacer las recomendaciones. Específicamente para el momento de construir las técnicas a utilizar. Por ejemplo, cuando se va a dar una recomendación puede ser útil saber qué tan parecido es un usuario con otro, o dos objetos

que se quieran recomendar, ya que de esta manera si a un usuario le gusta algún objeto en particular, podría esperarse que, a un usuario similar también, o si es que un usuario en particular tiene preferencia por algún objeto, también pueda tenerla con objetos parecidos. A este concepto se le llama similitud.

- Similitud: en el contexto de motores de recomendación, la similitud es una medida que permite comparar la proximidad o cercanía de dos objetos, de una forma muy parecida a por ejemplo la manera en que la distancia física entre dos edificios indica qué tan cerca están uno del otro geográficamente. La distancia y la similitud sólo difieren por el punto de referencia o por la coordenada espacial que está siendo usada. Hay que considerar también que por ejemplo dos ciudades pueden estar ubicadas de manera geográficamente similar, es decir, su distancia física es pequeña, pero culturalmente pueden ser muy diferentes, o sea su distancia, como medida de los intereses y hábitos de la población general, es muy alta. En general, el mismo cálculo de distancia puede ser realizado en diferentes espacios para obtener distintas medidas de similitud. [9]

Pensando en el mismo caso de las ciudades, se usarían sus coordenadas para calcular su proximidad geográfica, en el caso de los sistemas recomendadores lo que se utiliza es un rating dependiendo de los gustos de cada usuario, es decir, si consideramos el ejemplo de los inputs explícitos, se le pediría al usuario que califique una cierta cantidad de objetos y luego se compararía este vector con el de otro usuario para ver qué tan parecidos son, o sea qué tan grande es su similitud, o pequeña es su distancia.

Para el caso específico de la recomendación de preguntas PAES, no se le preguntaría al usuario por su calificación, ya que como se mencionó al momento de definir la dimensión interface y los tipos de inputs y outputs, este caso sería de inputs implícitos.

- Explicabilidad: es un concepto importante en los sistemas de recomendación ya que se da un trade off entre la calidad de la recomendación y su explicabilidad. Cuando se trabaja aplicando algoritmos para hacer mejores predicciones (en este caso recomendaciones), la exactitud de ellas se vuelve mejor, a medida que la interpretación de la misma se vuelve más compleja [10].

En el caso de las recomendaciones, mientras mejor sea su calidad, más complejo y difícil se vuelve la posibilidad de explicar el porqué de la recomendación. Este problema se conoce como *accuracy-model interpretation trade-off*. [11]

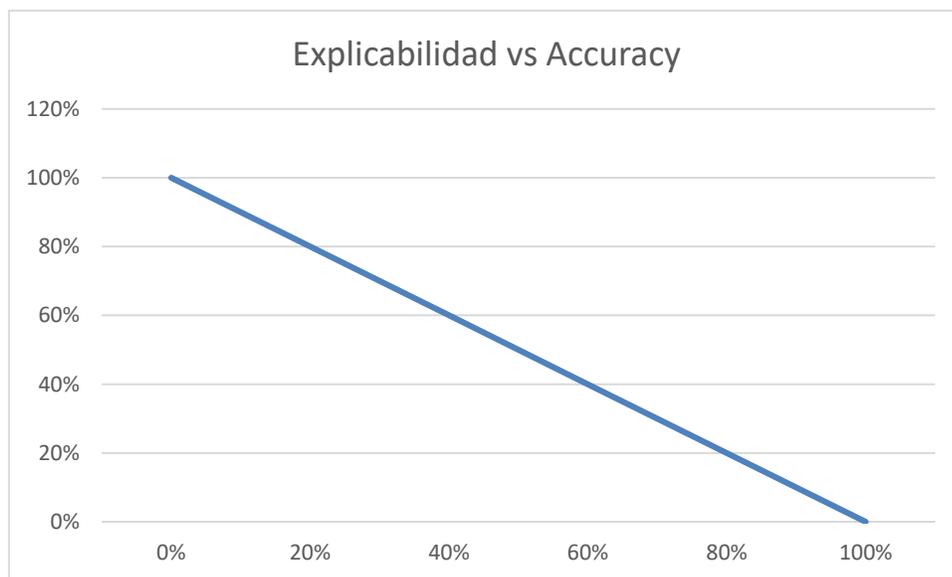


Figura 1: Explicabilidad Versus Accuracy, - Basado en "Accuracy-Model Interpretation Trade-Off" [11]

Con esto en mente, al momento de hacer el modelo y las técnicas de recomendación, hay que tener en cuenta un equilibrio para no llegar al extremo en que las recomendaciones sean tan difícil de entender y personalizadas que su implementación se vea entorpecida.

- Evidencia: es la data que revela el gusto del usuario [12]. Se le llama así porque es el "rastros" que nos deja el usuario para que se pueda inferir qué es lo que más le gusta hacer o los objetos que son de su preferencia.

En el caso de las recomendaciones en el ámbito del aprendizaje la evidencia ayudaría no a revelar los gustos o preferencias del usuario, sino que a revelar cuáles son los contenidos que se le hacen más dificultosos. Esto pasaría al recolectar evidencia de su comportamiento al hacerle preguntas que contestarán de acuerdo a lo que crean que es correcto, con esta información se podría deducir cuáles son los ámbitos que están más dominados por los usuarios en específico y cuáles no.

Es importante analizar con detenimiento la evidencia que se quiera recolectar y a qué hace referencia el sistema de recomendación en particular, porque como en el ejemplo de Netflix que ya se había mencionado la evidencia que se recolecta podría ser distinta a la que se haría en el caso de este trabajo.

En el caso de un sitio de venta, la evidencia que potencialmente se podría recolectar podría ser desde que el usuario comienza a buscar en la página, hasta que hacer el clic final antes de comprar [13]. Toda esta información se puede obtener, ya que en los sitios se puede ir guardando, creando sesiones, sin que necesariamente el usuario tenga una cuenta en el sistema. En el caso de este trabajo, no sería necesario tener que crear sesiones ya que cada usuario estaría previamente identificado, de manera de que si quiere detener su ejercitamiento podría continuarlo después dejándolo pendiente en su cuenta.

Comparando esta situación con la del sitio de ventas, habría evidencia que habría que evaluar si considerar o no al momento de crear la plataforma. Por ejemplo, antes de comprar el producto final en el caso de las ventas, puedo haber revisado otros de mi interés, lo que me ayuda a entregarle recomendaciones de ese tipo de artículos, el problema es que en el caso del sistema de recomendación de aprendizaje no es que el usuario esté más interesado en algún tipo de pregunta o no, ya que las recomendaciones aquí no se basarían en sus preferencias, sino que en su conocimiento y habilidad.

Estos indicadores se utilizan al momento de hacer las recomendaciones, se usan como parte de los factores a considerar del comportamiento de alguien al momento de recomendarle algo. Especialmente la similitud y la evidencia, por otro lado la explicabilidad es más bien una variable que entra en juego un poco después cuando se está creando el algoritmo que considerará distintas técnicas para realizar la recomendación.

2.1.4 Funciones de Similitud

Como se menciona en el punto anterior, la similitud es un factor muy importante al momento de empezar a entregar las recomendaciones, principalmente porque lo que se quiere es buscar objetos similares a los que al usuario ya le gustan, o encontrar personas que les guste lo mismo que le gusta al usuario. [14] Esta similitud se puede calcular de distintas maneras, pero en general se define como [15]:

Dados 2 items i y j , la similitud entre ellos está dada por la función $sim(i, j)$:

$$sim_{jaccard}(i, j) = \frac{\# \text{ usuarios que compraron } i \text{ y } j}{\# \text{ usuarios que compraron } i \text{ o } j}$$

La función devuelve un valor que crece a medida que la similitud entre ellos aumenta. Se podría decir que su similitud es $sim(i, j) = 1$ cuando i y j son el mismo item, y $sim(i, j) = 0$ cuando i y j no tienen nada en común.

Hay distintos métodos para medir similitud, y se recomiendan dependiendo del tipo de data que se tenga. A continuación, se describirán cuáles son recomendados dependiendo de la data [16]:

- Data Unitaria: esta es data que contiene sólo un tipo de objeto, podría ser por ejemplo solo *likes*, o sólo transacciones de ítems comprados. Ejemplo de esta data podría ser:
 - al usuario 1 le gusta la película 2.
 - al usuario 2 le gusta la película 1.

- al usuario 3 le gusta la película 2.

Para este tipo de data se recomienda utilizar la medida de similitud de Jaccard.

- Data Binaria: esta data es en la que se puede encontrar dos valores posibles. Por ejemplo, me gusta y no me gusta. Ejemplo de este tipo de data:

- al usuario 1 no le gusta la película 2
- al usuario 1 le gusta la película 1.
- al usuario 2 le gusta la película 1.

Para este tipo de data también se recomienda utilizar la medida de similitud de Jaccard.

- Data Cuantitativa: esta es la data que como su nombre lo indica, es cualitativa, es decir se puede medir en intervalos indicando que tanto te gusta o no algo, del tipo de preguntas califica un objeto de 1 a 10. Ejemplo de esta data:

- el usuario 1 le da 4 de 10 estrellas a la película 1.
- el usuario 2 le da 10 de 10 estrellas a la película 1.
- el usuario 3 le da 1 de 10 estrellas a la película 2.

Para este tipo de data se recomienda ocupar las medidas de similitud de Pearson o Cosine.

2.1.5 Medidas de Similitud

Como se acaba de describir en las funciones de similitud, dependiendo del tipo de data que se tenga es que se recomienda utilizar distintas medidas de similitud.

- Coeficiente de Similitud de Jaccard: esta distancia fue creada por Paul Jaccard para indicar que tan cerca estaban dos conjuntos uno del otro. También se puede encontrar como el índice de Jaccard o la distancia de Jaccard [17].

A simple vista dos conjuntos parecieran no tener relación con los tipos de recomendaciones que se necesitan, para un usuario y el contenido que se le desea recomendar. Es por eso que hay que pensarlo con una pequeña modificación, imaginando que cada película es un conjunto que contiene a todos los usuarios que calificaron positivamente a esa película, entonces de esa forma se pueden comparar las 2 películas, comparando los dos conjuntos de usuarios.

En el caso del sistema de recomendación de este trabajo se podría considerar cada conjunto como los usuarios que interactuaron con una pregunta en particular, es decir tendríamos un conjunto por cada pregunta. El conjunto podría ser definido como una lista de ceros y unos, donde los ceros indican si los usuarios respondieron la pregunta de manera correcta, y los unos si es que no acertaron a la respuesta correcta. Un ejemplo de conjuntos:

Preguntas	Usuario 1	Usuario 2	Usuario 3	Usuario 4	Usuario 5
Pregunta 1	0	1	1	0	1
Pregunta 2	1	1	1	1	0
Pregunta 3	1	0	0	1	0

Tabla 2: Ejemplo de Conjuntos - Elaboración Propia

En la tabla se puede ver que se tienen 3 conjuntos (cada fila indica el conjunto de la pregunta respectiva). Donde por ejemplo podemos ver que la pregunta 2 fue respondida correctamente por todos los usuarios, excepto por el usuario 5.

Para encontrar la similitud de Jaccard de dos preguntas hay que calcular cuantos usuarios respondieron incorrectamente ambas preguntas, y luego dividirlo por cuantos usuarios contestaron incorrectamente una o las dos preguntas [17]. Dado esto, la similitud

entre la pregunta 2 y 3 sería, la cantidad de usuarios que contestaron incorrectamente ambas preguntas (dos, el usuario 1 y el usuario 4), dividido la cantidad de usuarios que contestaron incorrectamente una o las 2 preguntas (cuatro, los usuarios 1, 2, 3 y 4). Por lo tanto, la similitud de las preguntas 2 y 3 sería $2/4 = 0,5$.

- Normas L: otra forma de medir distancias y similitudes son las normas. Si se tiene un ranking, donde el usuario evalúa ciertos objetos en alguna escala, por ejemplo, que el usuario 1 le ponga nota 5 a una cierta película, y que el usuario 2 le ponga nota 3, la distancia con la norma 1 se puede medir como:

$$\text{Distancia (usuario 1, usuario 2)} = |\text{ranking U1} - \text{ranking U2}|$$

Entonces la medida de similitud sería⁷:

$$\frac{1}{|\text{ranking U1} - \text{ranking U2}| + 1}$$

Donde mientras más se acerquen a 1 más parecidos son, mientras que el 0 los hace totalmente desiguales. [18]

La norma 2 tiene la misma idea, pero se define como la norma euclidiana:

$$\text{Distancia (U1, U2)} = \|\text{ranking U1} - \text{ranking U2}\|_2 = \sqrt{\sum_{i=1}^n |\text{ranking U1}, i - \text{ranking U2}, i|^2}$$

Cuya diferencia absoluta es:

$$\text{Suma de la Diferencia Absoluta} = \sum_{i=1}^n |\text{ranking U1}, i - \text{ranking U2}, i|$$

Y su diferencia medio absoluta:

⁷ Se le agrega este 1 para que no ocurra la división por 0, esto debe hacerse excepto si las 2 calificaciones son las mismas.

$$Diferencia\ Media\ Absoluta = \frac{1}{n} \sum_{i=1}^n |\text{ranking } U1,i - \text{ranking } U2,i|$$

En el caso del sistema de recomendación para preguntas PAES de matemáticas ambas normas no aplican ya que las preguntas tienen respuestas binarias (correcta o incorrecta).

- Similitud de Cosine: esta similitud considera otra forma de mirar el contenido, y es haciendo que la matriz de calificación de usuarios y los objetos que están calificando, pasan a ser vectores en el espacio, y lo que se observa para ver la similitud es el ángulo que se forma entre ellos. [19]

Se puede entender mejor usando un ejemplo. Considerando el plano de coordenadas, 2 objetos y supongamos 3 usuarios, se utiliza la calificación de cada usuario para ambos objetos, y dependiendo de estos valores es que la punta del vector tiene una ubicación en el eje x y en el eje y. Si se trazan los vectores desde el origen hasta estos puntos, se forman distintos ángulos entre los 3 usuarios, y estos son los que representan la distancia o también que tan similares son entre ellos.

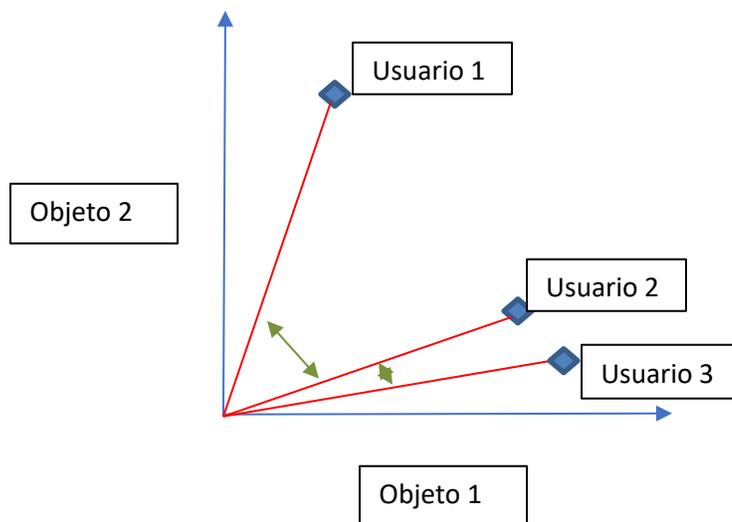


Figura 2: Esquema de ejemplo de Similitud de Cosine

Como puede verse en la figura 2, de acuerdo a esta medida el usuario 2 es más similar al usuario 3, que al usuario 1.

Al igual que el caso de las normas, para el caso de recomendación de preguntas PAES, es un poco más difícil ya que dado que los vectores podrían tomar solo los valores de 1 y 0, sólo quedarían en el eje x, en el y, o en exactamente 45 grados respecto a los ejes.

- Coeficiente de correlación de Pearson: el coeficiente de Pearson mide que tanto 2 líneas se parecen entre usuarios. Estas líneas se forman mediante el graficar la calificación que se le da a cada uno de los objetos. [20]

Por lo tanto, cada línea pertenece a un usuario y esta baja o sube si es que la pregunta esta contestada correcta o incorrectamente (ceros o unos respectivamente).

Cómo calcular el coeficiente de correlación de Pearson, a continuación:

$$sim(i, j) = \frac{\sum_{e \in U} (r_{i,u} - \bar{r}_i)(r_{j,u} - \bar{r}_j)}{\sqrt{\sum_{e \in U} (r_{i,u} - \bar{r}_i)^2} \sqrt{\sum_{e \in U} (r_{j,u} - \bar{r}_j)^2}}$$

Por ejemplo, tomemos a un usuario 1, si responde todas las preguntas correctamente su línea debería ser paralela y coincidente con el eje x, pues serían sólo ceros. Por otro lado, si es que lo respondió todo incorrecto, sería la misma línea paralela, pero con coeficiente de posición 1, si en un tercer caso algunas las responde correctas y otras incorrectas podría tener un zigzag entre ambos números que puede estacionarse o no en alguno de los 2 si es que coincide en correcto o incorrecto en más de una pregunta.

Pearson mira que tan diferentes son cada punto uno del otro, muy diferentes sería cercano a -1, mientras que muy parecido sería cercano a 1. Por otra parte, algo cercano a 0 significa que los ítems no estarían correlacionados.

2.1.6 Indicadores de Evaluación de Sistemas de Recomendación

Hay otro tipo de indicadores que se utilizan para medir el desempeño de un sistema de recomendación. Estos indicadores también se tendrán en cuenta al momento de que el sistema esté creado y se comience su implementación. Ellos son los indicadores de evaluación de sistemas de recomendación. A continuación, su definición y cómo sería su papel para el caso del sistema de recomendación de este trabajo en particular.

- **Efectividad:** se define una fase de aprendizaje en la que el usuario mostrará su preferencia por distintos objetos, la efectividad se mide considerando el porcentaje de objetos que se prefirió durante la fase de aprendizaje. [21] En general esto considera sistemas donde los usuarios acceden a contenido libremente, donde no hay una pauta a seguir al momento de responder las preguntas, como sería el caso del sistema de recomendación de este trabajo, donde luego de una pregunta viene otra que depende de lo respondido en las anteriores. Considerando esto, la efectividad podría medirse aquí como que porcentaje de las combinaciones posibles de eje temático – habilidad lograron superarse, o al menos mejorar en algún porcentaje. Esto habría que definirlo, podría ser que porcentaje de las combinaciones logró mejorar cierto porcentaje, que podría ser la meta de aprendizaje que se establezca con anterioridad.
- **Eficiencia:** indica el tiempo necesario para el usuario, para alcanzar su meta de aprendizaje. [21] Al igual que en el punto anterior donde hay que comenzar a incluir definiciones tales como que porcentaje de mejora será aceptable para calificarse como efectivo, aquí se tiene que establecer la meta de aprendizaje, es decir, cuanto es lo que se quiere mejorar en la totalidad de los ejes temáticos y habilidades que evalúa la PAES de matemáticas, y luego de tener este número, la eficiencia mediría cuantas iteraciones tomó poder llegar a esta meta.
- **Satisfacción:** es una medida subjetiva, que debe ser evaluada en conjunto con el usuario. [21] A pesar de que se tengan muchas

medidas que serán concretas y objetivas (con números), la percepción del usuario siempre es importante, es por eso que en este caso se podría considerar una encuesta final para todos quienes hicieron uso del sistema y que pudieran expresar si sintieron que aprendieron o que mejoraron más en algún ámbito en particular.

- Ratio de Abandono: representa el porcentaje de usuarios que dejan de participar en la configuración de aprendizaje, y que de esa forma finalmente no alcanzan la meta de aprendizaje establecida. [21] Al momento de realizar el experimento con usuarios reales para poder validar la hipótesis de investigación se espera que todos participen hasta el final, si esto no ocurriera se podría calcular este ratio de manera sencilla. Sin embargo, si es que todos los usuarios llegaron hasta el final se podría considerar el ratio de participación, en que momento decayó la participación, lo que se podría medir cuando esta pierde cierta frecuencia, que también tendría que definirse.

2.2 Plataformas E-Learning

El sistema de recomendación que se realizará en este trabajo es un sistema de recomendación de aprendizaje, por lo que es de vital importancia incluir también para su desarrollo la visión que se tiene desde las plataformas e-learning.

Los sistemas e-learning ya llevan bastante tiempo, y se utilizan en varios contextos de aprendizaje distintos. Con el crecimiento del uso de computadores, de internet, y de avances en las tecnologías de telecomunicaciones, el e-learning ha sido reconocido como una herramienta de aprendizaje muy valiosa. Es un tipo de aprendizaje a distancia donde el material para los estudiantes les es entregado de manera electrónica. [22]

Una plataforma e-learning provee de una infraestructura que integra material de aprendizaje, herramientas y servicios dentro de una misma solución para poder crear y entregar contenido educativo de manera rápida, efectiva y económica, por lo que se ha vuelto una alternativa relevante al aprendizaje

en las salas de clase. [22] Muchas veces en colegios y universidades son usadas también para que los alumnos puedan tener acceso a sus notas, o a comunicación con sus profesores y compañeros, pero en lo que a este trabajo concierne sólo será enfocado en la parte en la que se basa en contenido para el aprendizaje.

Las plataformas e-learning puede ser síncronas o asíncronas. Las síncronas requieren de la participación simultanea de todos los participantes, tantos de los que están para aprender como de quienes les van a enseñar, estando en diferentes lugares. Las plataformas e-learning síncronas serían cualquier tipo de evento en tiempo real de manera remota a usuarios que busquen aprender y además comunicación entre los participantes. Podría ser una video conferencia, una clase virtual o cualquier idea parecida. Por otra parte, las plataformas asíncronas no requieren de la participación simultanea de quienes enseñan y quienes aprenden, sino que hace referencia a un evento en donde el aprendizaje no ocurre en tiempo real [22]. Esto les da mayor libertad a quienes están buscando aprender, ya que pueden acceder a la plataforma en cualquier momento, y quedarse también tanto como deseen.

2.2.1 Características Plataformas E-Learning

Hay cuatro características importantes con las que las plataformas e-learning cumplen para poder entregarle una buena experiencia al usuario. Ellas son:

- Manejabilidad: este concepto hace referencia a poder hacer uso de la información que vaya surgiendo de la misma plataforma, desde saber quién está escribiendo o quién publicó algo, hasta poder registrar la actividad de cada usuario, por ejemplo, cuántos clic hizo antes de llegar a determinada página, o que fue lo que visitó antes. Todo este registro de actividad es especialmente valioso, ya que luego se podrían levantar indicadores con esta información y hacer mejoras.
- Accesibilidad: una de las características, quizás la más importante de las plataformas e-learning, es la accesibilidad. El usuario debe poder tener acceso, en todo momento y en cualquier lugar que tenga

conectividad, a la plataforma para poder revisar sus contenidos, sin necesitar estar en algún lugar determinado, sólo con los dispositivos para los cuáles la plataforma se encuentre habilitada para funcionar.

- Escalabilidad: las plataformas pueden ser escalables, idealmente deben serlo, de esta forma facilitan por ejemplo el agregar nuevo contenido para sus creadores, poder incluir nuevos cursos, o nuevas funcionalidades sin la necesidad de comenzar todo de nuevo.
- Optimización: esta es una característica inherente en todas las plataformas e-learning, por el sólo hecho de ser plataformas online. Hace referencia al hecho de eliminar acciones como:
 - Tener que asumir costos de viajar a algún lugar para poder tener el encuentro o el acceso a la información que se tendría en la plataforma.
 - Pérdida de tiempo en el viaje, permitiéndole tanto a quienes enseñan como a quienes buscan aprender, poder organizar mejor su tiempo.
 - Gastar en costos de producción de contenido online, esto se permite por la escalabilidad de las plataformas e-learning.

Actualmente los sistemas de e-learning pueden presentar algunos problemas. Es importante tenerlos en cuenta para evitarlos y así poder tener un mejor desempeño. Algunos de ellos son [22]:

- Material basado sólo en texto: uno de los principales problemas es que muchas veces presentan contenido sólo basado en lecturas, sin interacción, lo que hace que la experiencia se vuelva aburrida para el usuario, ya que generalmente son reacios a leer grandes volúmenes de texto [23], es por esto que resulta muy importante hacer que esto no ocurra y procurar la suficiente interacción.
- Falta de contenido con buenas explicaciones: un gran número de sistemas de e-learning no tienen explicaciones adecuadas para los estudiantes. Muchos de ellos sólo implementan diapositivas de power

point, y no es poco común que los estudiantes tengan dificultades y terminen por no entender el mensaje objetivo de la presentación [22]. En general el formato de las presentaciones es bien utilizado para presentar, pero esto se da porque el interlocutor maneja contenidos que no están totalmente detallados en las diapositivas, y aunque esto sea así, el complemento de ambos hace que se logre el objetivo. Sin embargo, cuando la presentación es utilizada en otro contexto como este, muchas veces no contienen el suficiente detalle, o al menos no en la forma que es necesario para poder entenderlo de manera óptima.

Por lo mencionado anteriormente es muy importante y relevante para que se cumpla con los objetivos de aprendizaje el no olvidar que se está en otro ambiente, en donde no se cuenta con las mismas herramientas que se tienen en un aula, por lo que hay que hacer que todo el material sea autocontenido y con las explicaciones necesarias para poder comprender a cabalidad el contenido presentado.

- Falta de flexibilidad: a diferencia de las salas de clases tradicionales en donde los estudiantes interactúan directamente con los profesores estando en el mismo lugar y en el mismo momento, puede llegar a ser un problema el hecho de tener que hacerlo de manera online, ya que requiere que los estudiantes hagan más esfuerzo. Esto, porque a veces el hecho de que los sitios no sean tan interactivos obliga a los estudiantes a estar durante todo el proceso, aunque por ejemplo sean cosas que ya sepan [22]. Debido a la poca flexibilidad que les es entregada puede que terminen perdiendo tiempo.

Esta característica no debería ser un problema para el sistema de recomendación de este trabajo ya que justamente se basa en lo que responde el usuario para capturar su aprendizaje y así continuar con las preguntas con las que el usuario tiene más problemas, de manera de tener siempre aprendizaje de algo nuevo, y luego reforzamiento de lo que presenta menor entendimiento.

- Instrucciones multimedia aisladas o poco estructuradas: en los últimos años la tecnología multimedia ha avanzado enormemente y ha tenido potencial para influenciar los procesos y productos ofrecidos por e-learning. A pesar de esto hay muchos sistemas de E-learning que simplemente publican contenido en la web sin hacer ningún procesamiento anterior. Las publicaciones son usualmente estáticas, pasivas, y sin estructura, además sin ningún tipo de asociación con otro contenido relevante, por ejemplo, para facilitar el aprendizaje, un video y una presentación de power point que mostrándose juntas se entienden mucho mejor, los estudiantes tienen que dirigirse a dos sitios distintos para poder revisarlas [22].

Dado esto la recomendación más importante aquí es tener esta característica en cuenta y preocuparse de que todo lo necesario esté en el mismo lugar, o para otro tipo de plataformas e-learning, como lo será el sistema recomendador de este trabajo, en el orden correcto.

Hasta el momento todo lo que se ha mencionado tiene que ver con plataformas e-learning genéricas, que no necesariamente implementan un sistema de recomendación. A continuación, la definición de lo que sería un sistema de recomendación para una plataforma de e-learning.

Un sistema de recomendación en un contexto de e-learning, es un agente de software que trata de recomendar acciones al aprendiz de manera inteligente, basándose en las acciones realizadas previamente por ellos. Estas recomendaciones podrían ser de actividades online como por ejemplo hacer un ejercicio, leer los mensajes publicados en un sistema de conferencia, correr una simulación online, o simplemente utilizar un recurso web. Estos sistemas de recomendación en su mayoría han tratado especialmente en el e-commerce de incentivar las compras de ciertos bienes, pero no han dedicado tanto tiempo al e-learning [24].

Como se menciona en [24], hay muchos estudios e investigaciones para sistemas de recomendación en e-commerce, pero cuando se trata del área de aprendizaje y educación todo se vuelve más difuso. Dado esto, lo que se

observará y tratará de estudiar son las plataformas e-learning en general y las que incluyen algún tipo de recomendación, en la mayoría de los casos en la forma de navegación. Desde aquí aparecen algunos conceptos nuevos. A continuación, su descripción:

2.2.2 Sistemas de Recomendación para E-Learning

Un "recomendador de tareas e-learning" es un sistema que recomienda una tarea de aprendizaje al usuario que está buscando aprender, basado en las tareas que el usuario ya ha realizado y en el éxito que ha tenido en ellas, además de las tareas realizadas por otros usuarios similares. La similitud entre los usuarios puede ser establecida usando perfiles de usuario, o puede ser basada en patrones de acceso previos. En principio hay 2 grandes partes en el diseño de este recomendador que son un módulo de aprendizaje, que aprende de patrones de acceso pasados e infiere un modelo de acceso común o individual, y un módulo recomendador, que aplica el modelo de aprendizaje en determinados momentos para así recomendar las siguientes acciones [24]. Hay muchas maneras de implementar estos procesos como por ejemplo con data clustering, asociación de reglas de minería o filtraje colaborativo [25].

Minando los datos de la web, particularmente los registros que se guardan en los servidores, el registro de ellos crea un rastro de la navegación de los usuarios en los sitios respectivos. Para poder procesar los datos de estos registros y obtener patrones de comportamiento para poder mejorar el sistema de aprendizaje, se necesita hacer una limpieza significativa y una transformación de los datos para preparar la información para los algoritmos de data mining que se van a utilizar [26]. Esto es principalmente considerando sitios de la web donde los usuarios disponen de distintas opciones y no de un portal guiado, y además muchas veces no se tiene registro de quien es el usuario que está interactuando por lo que es necesario, en base a los registros que se tienen, armar sesiones para poder definir un perfil y hacer recomendaciones en base a los patrones.

Sin embargo, en el caso del sistema de recomendación de este trabajo todos los usuarios tendrán una cuenta en la cual estarán registrados por lo cual no será necesario crear sesiones por medio del registro de los logs. Además, como el sistema de recomendación será de preguntas PAES mediante el aprendizaje de las respuestas a las preguntas anteriores, no hay navegación de forma desordenada, ya que los usuarios van contestando las preguntas que se les presentan, y ese es el camino que siguen.

2.3 Caso de éxito: Squirrel AI

El caso de Squirrel AI se parece bastante al problema que se aborda en este trabajo de tesis y es por eso que se le dedica este apartado. Squirrel AI es la primera empresa que se especializa en educación inteligente adaptativa en China, y es la líder del mercado. Fue fundada en el año 2014, y es el primer motor basado en el algoritmo avanzado desarrollado por YiXue Education. Se destaca por la experiencia que les provee a sus estudiantes, ayudándolos a avanzar en su aprendizaje a través de un sistema en tiempo real que se adapta y además incentiva hábitos de aprendizaje con práctica.

El sistema de AI simula a un profesor humano que le da al estudiante un plan de aprendizaje personalizado, dándole tutoría de uno a uno, resultando en tener de 5 a 10 veces más eficiencia que las instrucciones tradicionales. Squirrel AI ofrece cursos de alta calidad después del colegio tradicional, en ramos como Chino, Matemáticas, Inglés, Física y Química.

La plataforma de aprendizaje Squirrel AI le brinda a los estudiantes una experiencia supervisada de aprendizaje adaptativo, que ha sido probado que aumenta tanto la eficacia del estudiante como su compromiso con las plataformas de e-learning y los centros de aprendizaje presenciales que poseen. Hasta el momento, Squirrel AI ha abierto 1.700 colegios y tiene 3.000 profesores en más de 200 ciudades a lo largo de 20 provincias y regiones autónomas de China. Además, ha establecido un laboratorio de inteligencia artificial de educación en Nueva York y también colaboración como partners con algunas instituciones como la Universidad de Carnegie Mellon y UC-Berkeley. [27]

2.3.1 Análisis de Metodología de Squirrel AI

La metodología que utiliza Squirrel AI se basa principalmente en los computadores de los estudiantes. Los estudiantes y los profesores van a encontrarse a través de las pantallas. En una habitación están los estudiantes usando audífonos y teniendo una sesión en el computador, y en otra pueden haber tres estudiantes que están teniendo 3 clases de matemáticas de manera separada, junto a un profesor que los está monitoreando [28]. Como se puede observar son dos las modalidades que existen en el centro, con profesores presenciales que se dedican especialmente a un grupo reducido de alumnos para poder responderles de manera adecuada y personalizada, y también con tutores online.

Una de las mayores innovaciones que Squirrel AI ofrece es su escala y granularidad. Por cada curso que ofrece su equipo de ingeniería con un grupo de profesores máster subdividen la materia en una pieza conceptual lo más pequeña posibles. Por ejemplo, en el caso de matemáticas, la asignatura se puede dividir en 10.000 elementos atómicos, o "puntos de conocimiento", como por ejemplo números racionales, las propiedades de un triángulo o el teorema de Pitágoras. Lo que se busca con esto es diagnosticar las brechas de conocimiento en los estudiantes lo más precisamente que se pueda. Comparándolo con otro tipo de plataformas similares como ALEKS, una plataforma de aprendizaje adaptativo desarrollado por McGraw-Hill que fue la que inspiró a Squirrel AI, divide la misma materia en 3.000 puntos de conocimiento.

Una vez que los puntos de conocimiento son determinados, a los estudiantes se les asignan videos con cátedras, notas, ejemplos de trabajo y problemas prácticos. Las relaciones entre ellos, como tienen que ver uno con otro, y los temas que a veces tocan más de uno de estos puntos, se establecen en un gráfico de conocimiento, que también se basa en la experiencia de los profesores máster.

El estudiante comienza un curso de estudio con una prueba de diagnóstico para verificar que tan bien entiende los conceptos clave. Si responde

correctamente una respuesta que aparezca al principio el sistema asumirá que sabe los conceptos que están relacionados y los saltará, y con 10 preguntas, el sistema tendrá una idea grosso modo de lo que el estudiante necesita trabajar más y con eso le va a construir un currículo. A medida que el estudiante repase, el sistema actualiza el modelo de su entendimiento y ajusta su currículo de acuerdo a ello. Mientras más estudiantes utilicen el sistema se pueden ir viendo conexiones entre temas que no se habían visto antes, y que se van creando en base a lo que ellos van respondiendo. Mientras que la plataforma ALEKS hace también lo mismo, Squirrel dice que su optimización de machine learning es más limitada, lo que la haría menos efectiva en teoría. [28]

Al igual que la idea del sistema de recomendación de este trabajo, Squirrel AI busca ayudar a los estudiantes a obtener un mejor puntaje en sus pruebas estándar⁸ para poder entrar a la universidad. Y de la misma manera que la PAES de matemáticas de Chile, son pruebas con preguntas estandarizadas, lo que hace más fácil poder modelarlas.

2.4 Análisis Funcional de Plataformas Similares

Funcionalidades	Squirrel AI	Plataformas E-Learning
Ejercitación	Se pueden ejercitar distintas temáticas	Se pueden ejercitar distintas temáticas
Guía Personalizada	Si tiene, hay tutores que acompañan a los estudiantes en su aprendizaje	No tiene, el usuario es el que determina que pasos seguir
Costo	Costo elevado, ya que se debe financiar a los tutores personalizados	Bajo costo o en algunos casos gratuito
Acceso	Online y asistencia a sus laboratorios de enseñanza	Online

Tabla 3: Análisis de Funcionalidades - Elaboración Propia

⁸ Examen para entrar a las universidades en China, también puede darte acceso a universidades occidentales. Catalogado como uno de los más difíciles del mundo.

Las plataformas E-Learning dan solución parcial al problema ya que logran entregar herramientas para que los estudiantes puedan ejercitar las distintas temáticas, sin embargo, la solución no resuelve el problema en su totalidad, ya que, dado que el contenido está dispuesto de manera estándar, no existe una guía personalizada para que cada usuario pueda seguir el camino y dedicar el tiempo a cada materia que sea óptimo en su aprendizaje.

Por otra parte, está el Laboratorio Squirrel AI en China, que resuelve el problema, ya que cada estudiante cuenta con un plan personalizado con la guía de un tutor, pero requiere del tiempo de un tutor que está pendiente de un grupo de hasta tres estudiantes, por lo que nuevamente se escapa en cierta medida del alcance de muchos por su costo.

Capítulo 3: Levantamiento de Requerimientos

3.1 Levantamiento de *Features* Mínimas

Luego de haber revisado el marco teórico en lo que respecta al estado del arte de los sistemas de recomendación, las plataformas e-learning y los sistemas de aprendizaje adaptativos, esta sección hace referencia a un análisis y discusión de estos para poder hacer un levantamiento de las *features* mínimas que debería tener el sistema de recomendación.

3.1.1 Sistemas de Recomendación

Una de las dimensiones que describen los sistemas de recomendación es la interfaz. La interfaz describe los tipos de inputs y outputs que recibe y entrega el sistema. Este es un detalle importante a considerar para las *features* mínimas del sistema, ya que como se vio en la sección correspondiente (2.1.1), el input para el caso de este trabajo no puede ser explícito, por los mismos motivos mencionados, no se trata de clasificar algo que se está comprando dependiendo de los gustos del usuario, por el contrario, lo que se trata de inferir es la capacidad y el conocimiento que este maneja, que podría resultar difícil de contestar certeramente, o incluso con honestidad. Por lo que en este caso el input sería implícito basado en las respuestas que el usuario entregue a los ejercicios de matemáticas que se le vayan presentando. El output del sistema serían recomendaciones que respondan al conocimiento que el usuario va demostrando a medida que resuelve más ejercicios.

En cuanto a las técnicas de recomendación, estas serían colaborativas y basadas en contenido, dejando de lado las demográficas. Serían colaborativas, asumiendo que al igual que en las preferencias ante objetos en particular, también podría existir el hecho de tener habilidades similares basándose en el perfil de los usuarios. Considerando que, si son parecidos, las preguntas que se les propondrían podrían serlo también. Basadas en contenido pues dependiendo de su perfil, y de cómo se vaya comportando es

que se definirá que es lo que más le cuesta y esas serán las recomendaciones que se le realizarán. Se dejan totalmente de lado las recomendaciones demográficas, ya que para estos efectos no tiene relevancia ninguna de las características de demografía de los usuarios.

Respecto al enfoque con el cual se deberían hacer las recomendaciones, se aplicaría el *time based approach*, pero en una forma diferente a como es pensado originalmente, ya que en este caso a medida que avanza el tiempo y lo que tiene más relevancia es recomendar lo último, esto si sería así en caso de que la resolución sea con error, pero en caso de que la respuesta sea correcta, ésta no se mantiene correcta para siempre, sino que a medida que el tiempo va pasando se va considerando que se pierde el dominio (muy de a poco), pero no es algo que quede a perpetuidad, de esta forma si ha pasado mucho tiempo sin que se enfrente a una pregunta con esa temática será momento de recomendar otra vez para analizar si el dominio continua siendo el mismo, y si no lo es, dar más recomendaciones por ese lado para poder mejorar en las deficiencias.

Están también las reglas de asociación, que se tendrán en cuenta en caso de que sea necesario y se dé alguna relación de esa forma con alguno de los ítems. Si es que se observa algún patrón que indique que en la mayoría de los casos que hay error, también se tiene este mismo error en otro tipo de contenido será una observación que se tendrá en cuenta y se considerará incluir al momento de hacer las recomendaciones.

Entre otra de las *features* mínimas a incluir en el sistema de recomendación es considerar la evidencia que se va a recolectar de los usuarios, definir que va a ser lo importante en cada uno de los casos y poder analizar qué cosas se van a considerar al momento de poder armar el perfil de cada usuario. Una vez se tenga la evidencia recolectada y el perfil armado, en base a estos ítems hay que también incluir otra característica importante, la similitud. Como se revisó en la sección 2.1.4 hay varias formas de medir la similitud entre ítems o entre usuarios, dependiendo de con lo que se está recomendando y de qué tipo son las recomendaciones que se quieren realizar es que es el tipo de medida de similitud que se va a utilizar.

Probablemente lo más indicado sería utilizar la medida de similitud de Jaccard, esto, pensando en el tipo de objeto que se está recomendando. Las preguntas tienen sólo dos posibles outputs, correcto o incorrecto, dado que esta es una de las principales ventajas que tiene Jaccard versus otras medidas de similitud, es la que sería la más indicada.

3.1.2 Plataformas E-learning

En cuanto a las features mínimas que el sistema debería tener siendo una plataforma e-learning está su forma de entregar el contenido. Aquí las plataformas tenían la característica de ser síncrona (si se requería presencia simultánea de los participantes), o asíncrona si se podía utilizar la plataforma en cualquier momento, sin necesidad de que la contraparte estuviese también conectada, o también que quien quiera utilizar la plataforma lo tenga que hacer sujeto a ciertos horarios en los que el contenido estará disponible. La asincronía es muy relevante a considerar en este sistema, ya que eso le entrega más flexibilidad a la plataforma, no imponiendo restricciones para el momento de su uso.

El sistema debe tener también manejabilidad. La manejabilidad es el hecho de que sea sencillo acceder, manejar y trabajar con la información que se está generando. Esto es sumamente importante ya que se conecta directamente con una de las características mínimas que el sistema debía tener pensando en las *features* mínimas que poseen los sistemas de recomendación, esta característica es la evidencia. Para poder recolectar la evidencia es muy necesaria la manejabilidad, el poder tener un registro de la actividad del usuario, aquí probablemente se va a registrar que fue lo que visitó, cuánto tiempo pasó en cada pregunta, etc. También el usuario mismo, ya que cada usuario tendría una cuenta para llevar este registro, esto facilitaría el hecho de tener que crear sesiones basadas en los logs de las páginas web, como usualmente tiene que hacerse en sitios de compras donde los usuarios no están previamente registrados.

Otra característica importante con la que el sistema debe contar es con accesibilidad⁹, debe ser posible acceder a él desde todas partes en las que se cuente con internet, y además en cualquier momento que se desee. Es por esto que el sistema debe ser apto para manejarse desde computadores e idealmente también contar con una buena interfaz (que sea capaz de adaptarse) desde dispositivos más pequeños, como tablets o celulares, dando más facilidad para realizar la conexión.

El sistema debe poseer escalabilidad. Agregar nuevo contenido debe ser escalable, es decir, poder agregarlo de forma sencilla e independiente de la parte de la plataforma en donde se esté instaurando. Para el momento de agregar nuevas preguntas, estas simplemente deben ser registradas, y clasificadas de acuerdo a su eje temático y habilidad. De la mano de esto mismo viene la optimización como un agregado, ya que, gracias a la existencia del sistema, se ahorrarán los costos de movilización que en otro caso se habrían dado para ir a repasar presencialmente a algún lugar, además habría mejor administración del tiempo de los usuarios, ya que como se mencionó recién en cuanto a la movilización y sus costos, también se ahorra el tiempo de esta y tienen más tiempo para poder organizarse y administrarlo de mejor manera.

En cuanto a la forma en que el contenido es presentado a los usuarios del sistema, este debe ser dispuesto de manera que no existan largos texto que puedan confundir a los estudiantes, sino que existan buenas instrucciones, y se les dé el espacio a los usuarios para interactuar. Estas características mínimas, podría decirse que son inherentes al sistema ya que los ejercicios que se van presentando y las recomendaciones son realizados precisamente para crear interacción con el usuario.

Otra *feature* relevante es la flexibilidad que debe existir para que el usuario no se enfrente obligatoriamente a temáticas que el ya maneja. Esto al igual que las características mencionadas en el párrafo anterior, es algo inherente

⁹ No confundir con la accesibilidad para personas con alguna discapacidad física o cognitiva. Este punto hace referencia a que el sistema debe ser accesible desde todo tipo de dispositivos y desde cualquier parte que cuente con una conexión a internet.

al sistema (en términos de no darles ejercicios que ya saben al momento del estudio), ya que la plataforma se encarga de no hacerlo porque precisamente se dedica a repasar el contenido que está débil y no el que ya es sabido, debido a las recomendaciones que va entregando.

3.1.3 Squirrel AI

En el caso de las plataformas de aprendizaje adaptativo, como el caso de Squirrel AI, lo más destacable en su metodología para incluir en este sistema es la granularidad que se debe tener en cuanto a los distintos temas. Las preguntas de la PAES de matemáticas se dividen de acuerdo a su eje temático, y además de acuerdo a la habilidad que están midiendo. Los contenidos se dividen en 4 ejes temáticos, estos a su vez en 16 unidades temáticas, los que abarcan 45 subtemas, por otra parte, las habilidades que miden las preguntas pueden ser 4, por lo que hay un total de 180 combinaciones posibles para realizar preguntas.

3.2 Situación Actual

La situación actual considera el escenario planteado en los puntos 1.1 y 1.2 de esta tesis. El promedio de alumnos por sala de clases en Chile es superior a 30, lo que impide que los profesores puedan entregarle un grado de dedicación y personalización suficiente a cada niño, ya que las materias a revisar se encuentran establecidas en la programación de cada año y se debe avanzar aunque algunos estudiantes vayan quedando atrás, teniendo esto como consecuencia muchas veces el efecto bola de nieve que hace que dado que no se aprendió lo básico no se pueda aprender algo más complejo por no tener la base suficiente.

Luego, como alternativa a los profesores de colegio al que todos los niños tienen acceso, están las clases particulares, pero como se observó en el punto 1.2 en una conocida página de clases particulares chilena, los valores por hora de estas ascienden a valores entre \$15.000 y \$25.000 por hora, lo que hace infactible de poder tomar como alternativa permanente dado su alto costo, a una parte importante de la población que no podría incurrir en

este gasto. Si se consideran los quintiles en los que se basan las ayudas para estudiar en instituciones de educación superior, tres quintiles de la población no tendrían este "dinero de sobra" para poder apoyar a sus hijos, alrededor del 60% de la gente.

3.3 Diseño Conceptual

En base a lo estudiado en el marco conceptual y tomando en consideración las características que debería tener un sistema de recomendación educativo, la solución debe considerar input implícito en vez de explícito y también en cuanto a la cantidad de temas presentes para armar el sistema, este debe ser lo más granular posible, para así alcanzar el mayor nivel posible de especificidad y personalización para cada usuario.

Es por esto que lo más importante que debe tenerse en consideración para el diseño conceptual de la solución es:

- Sistema Online: la solución debe estar disponible de manera online ya que esto aumenta la accesibilidad para los usuarios y hace que puedan utilizarlo con mayor facilidad
- Ejercitación: no debe ser una plataforma que solo entregue contenido que haga que los usuarios tengan un comportamiento pasivo (sólo leer), sino que debe proporcionar la característica de interacción, donde los usuarios puedan ejercitar.
- Guía Personalizada: la ejercitación que cada usuario realice debe ser con una guía, cada usuario debe tener un plan personalizado dependiendo de su desempeño/conocimiento en las distintas temáticas que se presenten.
- Costo: no debe tener un costo alto como el que tendrían las clases particulares o el del pago de tutores en un aula común, debe ser un costo bajo (por ejemplo, el costo mensual podría análogo a lo que costaría una hora de clases en un aula común) para que el costo no sea un impedimento para que ningún estudiante pueda utilizarlo.

Capítulo 4: Propuesta de Solución

4.1 Propuesta de Solución

Considerando los puntos más importantes levantados que debe tener el diseño conceptual, la solución al problema presentado, “la falta de estudio/ejercitación/guía en los estudiantes los lleva a un fracaso académico” [29], es un sistema online de recomendación de preguntas PAES de Matemáticas, que esté al alcance de todos, y como su nombre lo indica pueda realizar recomendaciones a cada estudiante de manera personalizada para que así cada uno pueda ir mejorando en los temas que particularmente tiene más débiles.

Lo que se va a implementar será una plataforma online en la cual los distintos usuarios puedan crearse una cuenta, allí, realizarán una prueba de diagnóstico para que el sistema pueda evaluar cuál es su nivel en las distintas temáticas que abarca la prueba y en base a eso le comience a entregar preguntas para que pueda comenzar a ejercitar, y a medida que siga respondiendo preguntas el sistema le recomiende las siguientes en base a lo que él ha respondido, es decir que le recomiende preguntas de las temáticas en las que sigue cometiendo errores.

Lo que se está mejorando en este proceso es el tipo de preguntas que el usuario va respondiendo y el orden en el que se presentan, ya que no sería de cualquier temática, sino que de las temáticas en que el usuario tiene un nivel más bajo, y así mismo, el orden de las preguntas busca en cada nueva iteración enfocarse en donde el sistema evalúa que el usuario tiene el menor conocimiento.

Esta tesis tributa al Magister en Gestión de Operaciones ya que el algoritmo que está detrás del sistema recomendador de preguntas está optimizando lo que haría un sistema tradicional, entregar preguntas al azar, o dejarlo a criterio del mismo usuario. En este caso se está constantemente revisando cuál es la temática en la que el usuario necesita seguir ejercitándose, en

base a lo que el mismo usuario ha entregado como input al sistema, así se optimiza el aprendizaje, ya que en teoría (y lo que se busca comprobar con la hipótesis de investigación), los usuarios que utilicen este sistema deberían tener mejor desempeño que los estudiantes que utilicen un sistema aleatorio, o en su defecto obtener el mismo desempeño, pero con menor número de iteraciones.

4.2 Modelo de Proceso

A continuación, el listado de etapas del proceso que presenta la solución, Kuizus, el Sistema de Recomendación de Preguntas PAES de matemáticas, con su descripción y luego el modelo de procesos representado gráficamente para un mejor entendimiento de cómo va avanzando en proceso y qué caminos puede tomar.

4.2.1 Usuario Ingresa a la Plataforma

Aquí el usuario accede al sitio web de la plataforma, y dependiendo de su estado, el siguiente paso que puede seguir es iniciar sesión, donde ingresa su usuario y contraseña, o si no se ha creado una cuenta aún, se dirige a registro, para poder crearla.

4.2.2 Usuario Inicia Sesión

El usuario está en la pantalla de inicio de sesión. Aquí introduce su usuario y contraseña creados al momento de registrarse, y enviando estos datos puede ingresar a su cuenta.

4.2.3 Usuario Se Registra

El usuario está en la pantalla de registro. Aquí, para crear una cuenta, ingresa su nombre, una contraseña y su correo electrónico para poder verificar su cuenta.

4.2.4 Usuario Verifica su Cuenta

El usuario accede a su correo electrónico y abre el correo enviado por Kuizus para darle clic al link de confirmación y con esto terminar el proceso de verificar su cuenta. Al hacer esto su cuenta queda verificada y puede comenzar a utilizar la plataforma.

4.2.5 Inicio de la Plataforma

En la pantalla de inicio aparecerá una breve descripción de lo que se puede hacer en este sistema de recomendación para darle contexto al estudiante y un botón para que pueda comenzar con sus preguntas. El botón dirá "Comenzar a ejercitar".

4.2.6 Comenzar a Ejercitar con Diagnóstico

Aquí se desplegará un mensaje que le comunicará a los usuarios que para poder comenzar a ejercitar primero se les realizará una prueba de diagnóstico que conste de 20 preguntas, y que con ello el sistema podrá tener una línea base para saber que preguntas comenzar a recomendarle.

4.2.7 Ejercitación

A medida que se vaya avanzando en responder preguntas el sistema va a tener más información para poder continuar haciendo más recomendaciones y así también poder ir actualizando los datos que haya de cada combinación posible (habilidad-eje temático).

4.2.8 Usuario Está Ejercitando

En cada momento en que el usuario está ejercitando, se le mostrará la pregunta, las alternativas, la opción de marcar una alternativa, de seguir más tarde, de pasar a la siguiente pregunta, o de volver al inicio.

4.2.9 Pantalla de Inicio

Cuando el usuario vuelve a la pantalla de inicio una vez que ya realizó la prueba de diagnóstico y ha comenzado a responder preguntas después de ello, puede acceder a una pantalla que le muestra que porcentaje de preguntas ha respondido bien, y que porcentaje mal, así puede hacerse una idea de cómo le está yendo y puede ir revisando su evolución.

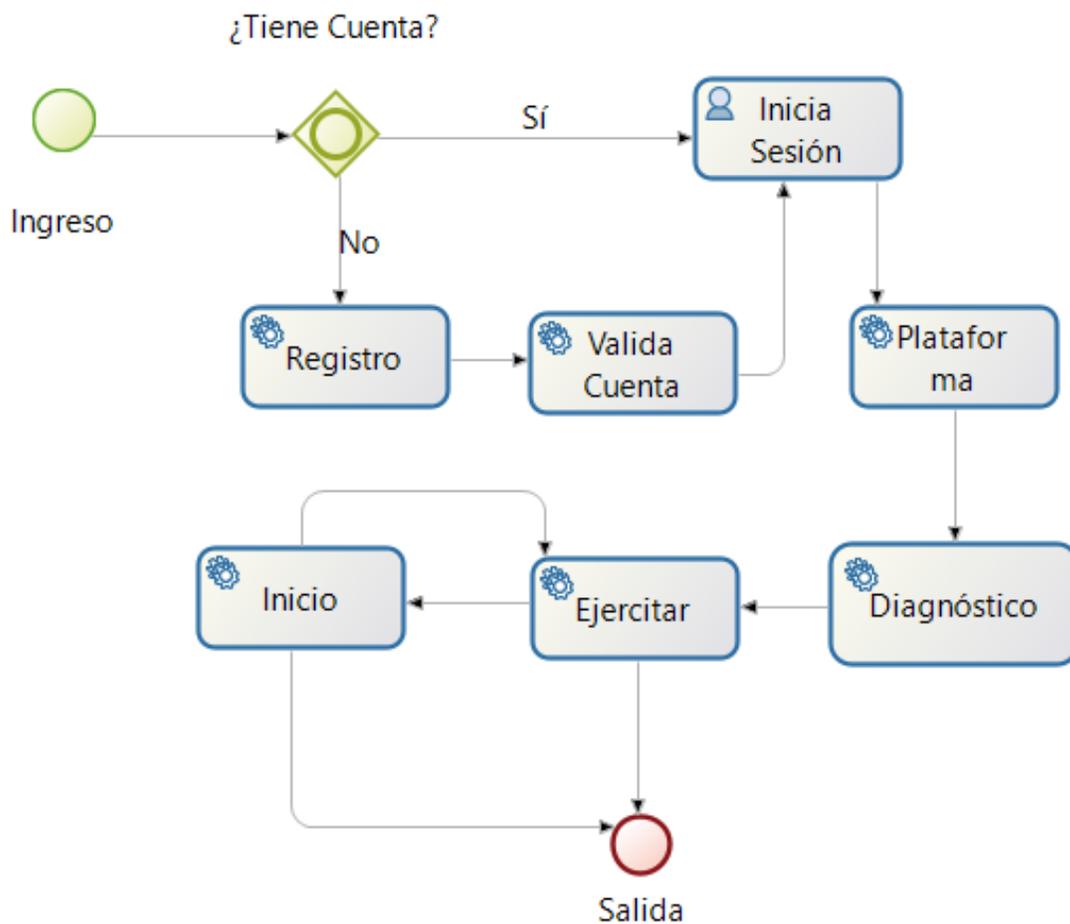


Figura 3: Proceso del Usuario en la Plataforma - Elaboración Propia

4.3 Categorización y Obtención de Preguntas PAES

La prueba PAES de matemáticas tiene varios tipos de preguntas, éstas se clasifican dependiendo de los ejes temáticos que tienen y las habilidades que buscan medir. Los contenidos a medir en esta prueba están agrupados en los siguientes Ejes Temáticos¹⁰:

- Números
- Álgebra y Funciones
- Geometría
- Probabilidad y Estadística

Luego, de acuerdo a los objetivos fundamentales por eje, se dividen en 45 subtemas que pueden verse en el anexo A.

Al hacer esta nueva clasificación, quedan los 45 “tags” finales. Además, de acuerdo con el temario PAES de matemáticas se definen cuatro criterios para definir las habilidades que se están midiendo en las preguntas y éstas son:

- Resolver problemas
- Modelar
- Representar
- Argumentar

La descripción [30] de cada una a continuación:

- Resolver Problemas: es la habilidad que se tiene para solucionar una situación problemática dada, contextualizada o no, rutinaria o no, con o sin que se le haya indicado un procedimiento a seguir. Para ello, puede realizar cálculos, aplicar diversos conocimientos y estrategias, además de interpretar y evaluar sus resultados, a través del pensamiento reflexivo, crítico y creativo.

¹⁰ estos ejes temáticos provienen del temario oficial de la prueba.

- Modelar: es la habilidad que se tiene para usar, entender y comparar expresiones matemáticas que describen las características relevantes de una situación de la vida diaria o de las ciencias, para poder estudiarla, entenderla y obtener soluciones en el ámbito matemático que permitan responder preguntas acerca de la situación modelada.
- Representar: es la habilidad que se tiene para transferir información de situaciones diversas que involucren datos, procesos y relaciones, mediante el uso de símbolos propios del lenguaje matemático, utilizando tablas, gráficos, diagramas, la recta numérica o el plano cartesiano y transitar entre las diferentes representaciones.
- Argumentar: es la habilidad que se tiene tanto para reconocer, explicar y justificar la validez de un procedimiento, de pasos deductivos o de una demostración, de estrategias de solución de diversos problemas, detectar argumentos erróneos en proposiciones del tipo "si se tiene esto, entonces se cumplirá esto otro", como por ejemplo, al invertirla.

4.4 Base para las Recomendaciones

4.4.1 Input Implícito

El algoritmo se basará en el input implícito que ingresan los usuarios mediante las respuestas que dan a los ejercicios que se les van presentando. Como se mencionó anteriormente en el marco teórico, no es aconsejable utilizar el input explícito en ámbitos de sistemas de recomendación educacionales, ya que puede levantarse falsa información preguntándole a los usuarios que es lo que ellos creen, y lo más cercano a la realidad es que se sometan a contestar preguntas de los temas en cuestionamiento, y así se puede saber con mayor certeza cuáles son los que manejan y los que no, basándose en si sus respuestas fueron correctas o erradas.

4.4.2 Asociación de Items

Además de lo recién mencionado se considerará también la asociación de ítems para incorporarla a las recomendaciones, así si más de un número determinado de usuarios tiene un mismo par o grupo de combinaciones incorrectas (una combinación se determina por habilidad y contenido), a los próximos usuarios y a los mismos involucrados se les recomendaría la otra combinación si es que falla en una de ellas.

Un ejemplo de la asociación de ítems para poder entender mejor la idea:

- Usuario 1: contesta mal pregunta contenido5 habilidad2 y contenido3 habilidad2.
- Usuario 2: contesta mal pregunta contenido3 habilidad2, contenido5 habilidad1 y contenido5 habilidad2.
- Usuario 3: contesta mal pregunta contenido11 habilidad1, contenido5 habilidad2 y contenido3 habilidad2.

Como se puede observar en el caso de los tres usuarios se repiten las dos combinaciones contenido5 habilidad2 y contenido3 habilidad2, por lo que por la asociación de ítems que se está incorporando, en futuras ocasiones si alguno de estos 3 usuarios, o cualquier otro responde incorrectamente alguna pregunta de la combinación contenido5 habilidad2, se le sugerirá una de la combinación contenido3 habilidad2 y viceversa, ya que de acuerdo a la asociación de ítems, si tiene mala una pregunta de una de estas dos combinaciones, es probable que también necesite ejercitar la otra.

Esto puede calcularse con la medida de similitud de Jaccard, creando conjuntos de usuarios para cada una de las preguntas, de esta manera se puede ver la similitud que existe entre las preguntas observando los conjuntos de usuarios de cada par de preguntas.

4.4.3 Similitud entre Usuarios

Al igual que en el caso de la asociación de ítems, si es que dos usuarios se parecen entre sí, esto será considerado para las recomendaciones. Es decir,

si dos usuarios tienen varias combinaciones similares, probablemente cercano al 50% de las combinaciones que cada uno de ellos ha respondido es un indicador de que se parecen. Para medir esta similitud entre usuarios también se puede usar la medida de similitud de Jaccard, esta vez los conjuntos serían las preguntas respondidas por cada uno de los usuarios, y se calcularía la similitud entre los conjuntos para ver qué tan similares son los usuarios.

4.5 Algoritmo Recomendador

El algoritmo recomendador se basa en el concepto de time based approach, pérdida de memoria y en la medida de similitud de Jaccard, revisados en el Marco Teórico.

Los usuarios comienzan a contestar preguntas, si responden correctamente, el puntaje actual que tiene en esa combinación se actualiza sumándole uno. Si responde de manera errónea, la combinación respectiva se actualizará restándole uno.

El usuario recibe 20 preguntas aleatorias que armarán el diagnóstico inicial (con esto el usuario tendrá puntaje en 20 combinaciones distintas). Una vez finalizadas las 20 preguntas, comienzan las recomendaciones. De todos los puntajes de las distintas combinaciones, el algoritmo seleccionará la combinación que tiene menor puntaje y de esa combinación le realizará la siguiente pregunta, si hay algún empate en puntajes, se elegirá la combinación de manera aleatoria.

Si el usuario responde correctamente esa pregunta se actualiza el valor de esa combinación sumándole 1, si responde de manera errónea se le restará 1. En cada iteración desde la pregunta número 21 (que ya no forma parte del diagnóstico y será la primera recomendación basada en el input del usuario) se multiplicarán todos los puntajes de las distintas combinaciones para cada usuario, por 0.99 reflejando así, la pérdida de memoria. Este es el valor que se utiliza por defecto, pero es un parámetro ajustable.

4.5.1 Pseudo Código del Algoritmo

A continuación, el detalle del algoritmo paso a paso para un usuario en particular:¹¹

- Sea A la matriz $(a_{1,1} \ a_{1,2} \ \dots \ a_{i,j} \ \dots \ a_{3,42} \ a_{3,43})$, $i \in \text{habilidades}(1, 2, 3)$ y $j \in \text{contenidos}(1, 2, \dots, 42, 43)$.
- Sea P el vector de preguntas $(p_{1,1,1,l} \ p_{1,1,2,l} \ \dots \ p_{i,j,k,l} \ \dots \ p_{3,43,n-1,l} \ p_{3,43,n,l})$, $i \in \text{habilidades}(1, 2, 3)$, $j \in \text{contenidos}(1, 2, \dots, 42, 43)$, $k \in \text{preguntas}(1, \dots, n)$ y $l \in \text{alternativas}(1, 2, 3, 4, 5)$
- Sea D el vector de preguntas de diagnóstico $(d_1 \ d_1 \ \dots \ d_i \ \dots \ d_{19} \ d_{20})$, $i \in \text{preguntas}$
- Sea $r_{k,l}$ el conjunto de alternativas de una pregunta p_k , $k \in \text{preguntas}(1, \dots, n)$, y $l \in \text{alternativas}(1, 2, 3, 4, 5)$, $\exists ! l \text{ correcto } \forall k$

1. Inicio
2. Se inicializa A en 0.
3. Usuario responde D , $\forall d \in D$ se guarda $r_{k,l}$,
4. Para $r_{k,l}$ desde $k = 1$ hasta $k = 20$ y $l \in \text{alternativas}(1, 2, 3, 4, 5)$
 - si $l = \text{correcto}$
$$a_{i,j} \leftarrow (a_{i,j} + 1)$$
 - si $l = \text{incorrecto}$
$$a_{i,j} \leftarrow (a_{i,j} - 1)$$
5. Actualizar $a_{i,j}$, encontrar menor $a_{i,j}$, " $\check{a}_{i,j}$ " y entrega al usuario $\check{p}_{i,j,k,l}$
 - si se recibe $l = \text{correcto}$
$$a_{i,j} \leftarrow (a_{i,j} + 1)$$
 - si se recibe $l = \text{incorrecto}$
$$a_{i,j} \leftarrow (a_{i,j} - 1)$$
$$\forall a_{i,j} \leftarrow a_{i,j} * 0,99$$
6. Repetir 5 hasta que usuario deje de responder preguntas
7. Fin

4.5.2 Ejemplo de Funcionamiento del Algoritmo

¹¹ Para detalle gráfico del proceso revisar Anexo B

La prueba de diagnóstico inicial consta de preguntas con las distintas combinaciones que pueden hacerse entre habilidades y contenidos, los contenidos son 43 y las habilidades 3, por lo que las combinaciones posibles podrían ser 129. Hacer una prueba de diagnóstico de 129 preguntas puede ser abrumador para un diagnóstico, por lo que este contará solamente con 20¹² preguntas aleatorias de distintas combinaciones para poder luego de ello comenzar con las recomendaciones, y a medida que estas vayan surgiendo ir incluyendo también el resto de las combinaciones posibles para poder ir acumulando información de estas también.

Cada vez que se contesta una pregunta se guarda la respuesta en la combinación habilidad-contenido respectiva para cada uno de los usuarios. Cada usuario tendrá una tabla con las distintas combinaciones partiendo con un valor de 0 en ellas, y a medida que vaya respondiendo preguntas se les asignará puntaje. En caso de que la pregunta sea respondida correctamente se le sumará un punto a la combinación respectiva, en caso de que sea respondida de forma incorrecta, se le restará un punto. Además, con el paso del tiempo por cada iteración nueva que se realice (avanzar con las preguntas), se les multiplicará a los valores que tengan los puntajes de las combinaciones, por un factor menor a 1, pero cercano a este, para mostrar cómo según el concepto de time based approach a medida que el tiempo avanza las últimas respuestas correctas son las que más valor tienen, y a medida que el tiempo avanza éste va disminuyendo.

De esta forma se irán obteniendo distintos valores para las distintas combinaciones que existen, y el algoritmo en el que se basará la recomendación primeramente será en hacer que la siguiente pregunta que deba responder sea la que esté más baja en puntaje. Un ejemplo de parte de esta tabla sería:

Usuario	Habilidad	Contenido	Combinación	Puntaje
Usuario 3	Habilidad 1	Contenido 12	1-12	1
Usuario 3	Habilidad 1	Contenido 3	1-3	1

¹² se consideran 20 preguntas, ya que según lo que se pudo ver en el marco teórico este es el mínimo necesario de ítems que el usuario debiese calificar para comenzar a entregar recomendaciones

Usuario 3	Habilidad 3	Contenido 3	3-3	0
Usuario 3	Habilidad 2	Contenido 40	2-40	0
Usuario 3	Habilidad 2	Contenido 35	2-35	1
Usuario 3	Habilidad 1	Contenido 5	1-5	-1
Usuario 1	Habilidad 3	Contenido 3	3-3	1
Usuario 1	Habilidad 3	Contenido 11	3-11	-1
Usuario 1	Habilidad 2	Contenido 9	2-9	0
Usuario 1	Habilidad 1	Contenido 7	1-7	0
Usuario 4	Habilidad 3	Contenido 5	3-5	0
Usuario 4	Habilidad 2	Contenido 21	2-21	1
Usuario 4	Habilidad 3	Contenido 5	3-5	1
Usuario 2	Habilidad 2	Contenido 17	2-17	-1
Usuario 2	Habilidad 3	Contenido 29	3-29	1

Tabla 4: 1° iteración de tabla de puntajes de los usuarios – Elaboración Propia

En la tabla pueden verse las distintas combinaciones de habilidad – contenido para 4 usuarios distintos. Es importante notar que la tabla es acumulativa, por lo que no se puede repetir la misma combinación para un usuario. Este ejemplo muestra posibles valores que dejaría la prueba de diagnóstico realizada a los usuarios.

Respecto a los números que aparecen en la columna puntaje, los valores que dicen 1 en puntaje significan que se les dio un punto por haber respondido correctamente una pregunta de la combinación respectiva, en el caso del 0 es porque aún no se han enfrentado a ninguna pregunta de esa combinación, y por último los que tienen el valor -1 significa que contestaron la pregunta de manera incorrecta.

Ahora, para ver cómo funciona el descuento por el concepto de time based approach, en la siguiente tabla se muestra la primera iteración realizada luego de la prueba de diagnóstico, donde la pregunta que se realiza a cada uno de los usuarios es la que indica la recomendación, es decir la combinación 3-11 para el usuario 1, 2-17 para el usuario 2, 1-5 para el usuario 3 y 3-5 para el usuario 4.

Usuario	Habilidad	Contenido	Combinación	Puntaje
Usuario 3	Habilidad 1	Contenido 12	1-12	0.99
Usuario 3	Habilidad 1	Contenido 3	1-3	0.99
Usuario 3	Habilidad 3	Contenido 3	3-3	0
Usuario 3	Habilidad 2	Contenido 40	2-40	0
Usuario 3	Habilidad 2	Contenido 35	2-35	0.99
Usuario 3	Habilidad 1	Contenido 5	1-5	0
Usuario 1	Habilidad 3	Contenido 3	3-3	0.99
Usuario 1	Habilidad 3	Contenido 11	3-11	-1
Usuario 1	Habilidad 2	Contenido 9	2-9	0
Usuario 1	Habilidad 1	Contenido 7	1-7	0
Usuario 4	Habilidad 3	Contenido 5	3-5	1
Usuario 4	Habilidad 2	Contenido 21	2-21	0.99
Usuario 4	Habilidad 3	Contenido 5	3-5	0.99
Usuario 2	Habilidad 2	Contenido 17	2-17	-2
Usuario 2	Habilidad 3	Contenido 29	3-29	0.99

Tabla 5: 2° iteración de tabla de puntajes de los usuarios – Elaboración Propia

Se puede observar que el usuario 1 tuvo mala la pregunta que se le recomendó por lo que el valor disminuye a -1, en el caso del usuario 2, la respondió mal de nuevo, así que queda en -2, en el caso del usuario 3 la contesta de forma correcta, por lo que se le suma 1 y queda en 0 y en el caso del usuario 4 ya estaba en 0 y queda en 1 gracias a haber respondido correctamente. El concepto de time based approach se puede observar en el resto de las preguntas que tenían sobre 0 (las que habían respondido bien y tenían el valor de 1) y que fueron multiplicados por 0,99 para así disminuir un 1% su memoria.

Además, se incorporarán para su análisis evidencia como, la duración del tiempo que el usuario pasa en alguna pregunta, es decir, cuánto se demora el usuario en responder la pregunta. También las alternativas que marcó antes de enviar la respuesta definitiva.

Para generar recomendación de similitud entre pares, en cada momento que se vaya a generar la siguiente pregunta para hacer la recomendación se va a

buscar si hay alguien parecido al usuario, o sea que tenga al menos 2 combinaciones similares (en cuanto a lo positivo). Y con esto se va a buscar al que más combinaciones similares tenga, una vez seleccionado, la combinación más negativa que tenga será la que se le recomendará al primer usuario. Esta similitud se calcula mediante la medida de Jaccard, y si tienen cerca del 50% de coincidencia es porque se parecen.

4.6 Arquitectura a Implementar

Para poder hacer el sistema de recomendación, la arquitectura a implementar se dividirá en tres partes, estas serían la aplicación web, el motor recomendador y la base de datos.

4.6.1 Aplicación Web

La aplicación web es el lugar donde la persona interactúa con el sistema recomendador en la página web. Dado que la interacción que el usuario realiza con la página es acotada, no se requiere de nada muy sofisticado, es por esto que las tecnologías estándar sirven para la aplicación, ya que el rol que tiene la aplicación web es ser capaz de desplegar preguntas para mostrar al usuario (preguntas que serán seleccionadas por el motor de recomendación), y luego que el usuario las responda, guardar las respuestas de las mismas, que el usuario ingresó como input, para que luego el motor recomendador pueda obtener información de estas para poder seguir realizando recomendaciones.

4.6.2 Motor Recomendador

El motor recomendador es el encargado de interpretar las respuestas que los usuarios dan a las preguntas que les son entregadas por la aplicación web, de manera tal de poder ir guardando los nuevos resultados, e ir así actualizando los resultados del usuario para entregarle nuevas recomendaciones de acuerdo al input que se ingresa.

Cuando llegue una solicitud de pregunta, esa petición la recibe la aplicación web, pero la envía directamente al motor recomendador, el motor la procesa y le entrega una respuesta a la aplicación web, que luego la aplicación web le entrega al usuario.

Para poder hacer que el motor recomendador pueda comunicarse con la aplicación web se puede utilizar un web service para que los conecte. Esta fue la forma que se utilizó en el desarrollo de esta tesis.

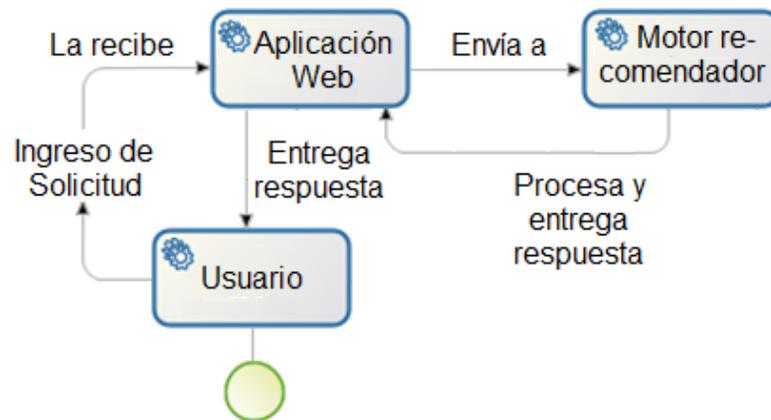


Figura 4: Proceso de Interacción Aplicación Web y Motor Recomendador - Elaboración Propia

4.6.3 Base de Datos

La base de datos es el lugar donde se almacena todo, tanto las preguntas posibles para realizar a los usuarios, cómo las respuestas que dan de ellas, y a los mismos usuarios.

La aplicación web se conecta a la base de datos para poder leer las preguntas, también para poder almacenar las respuestas y para acceder a los usuarios que posee el sistema. Tanto la aplicación web como el motor recomendador pueden conectarse a la base de datos sin necesidad de un web service u otro elemento. Sin embargo, lo que ven la aplicación web y el motor recomendador no es necesariamente lo mismo (podría serlo). Lo que tiene que ver el motor recomendador es lo que permita generar las recomendaciones, y lo que ve la aplicación web es lo que permita dar y recibir preguntas.

4.7 Modelo de Datos

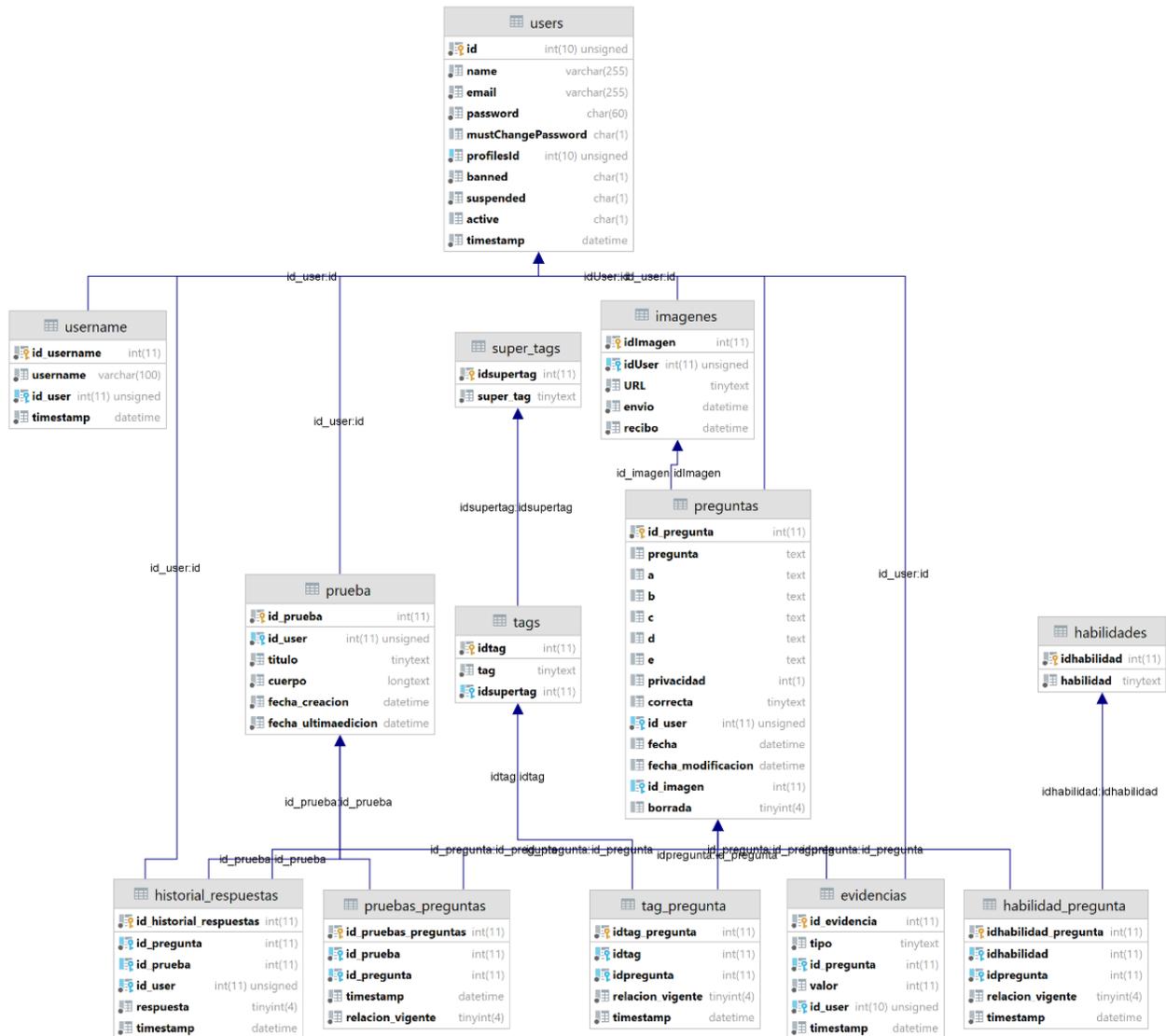


Figura 5: Diagrama Entidad Relación - Elaboración Propia

A continuación, se explicarán las tablas más relevantes del modelo de datos:

- **Preguntas:** esta tabla almacena todas las preguntas que están en el sistema. Se modeló pensando en que siempre se tendrán preguntas estandarizadas con 5 alternativas. También considera si es que la pregunta está activa, y si es que tiene alguna imagen asociada.

- Tags y Habilidades: estas dos tablas son las que tienen las características especiales de cada pregunta. Se decidió separarlas para tener un orden sobre estas. Tag corresponde al contenido específico de la materia relacionada a la pregunta. Por otro lado, se tiene la tabla Habilidad, la que contiene cada una de las habilidades que mide una pregunta.
- Historial Respuestas: esta tabla es una de las más relevantes para el modelo recomendador, ya que a partir de ésta nace la primera recomendación que se hace, y que luego va evolucionando con el tiempo.
- Evidencias: esta tabla tiene por objetivo, el de guardar todas las evidencias que se pueden registrar, asociadas a lo que se puede capturar del proceso de respuesta de una persona, como, por ejemplo, el tiempo que le tomó responder una pregunta. Como se ha dicho anteriormente, esto se considera en el modelo recomendador, y servirá para ir robusteciéndolo a futuro.

4.8 Modelo de Interfaz

En esta sección se verán las vistas más relevantes para este trabajo. Comenzando con las preguntas del diagnóstico, para luego seguir con la vista de entrenamiento (que va mostrando las preguntas recomendadas).

4.8.1 Vista de Diagnóstico

La vista de diagnóstico muestra la cantidad de preguntas que tiene el diagnóstico total, la que se almacena internamente en el sistema como una prueba. Los usuarios aquí no pueden saber si las respuestas que enviaron son o no correctas, sin embargo, pueden cambiar de alternativa si lo desean, incluso después de apretar "responder". Todos los movimientos son almacenados como evidencias para poder tener un registro de lo que ocurre durante el proceso de respuesta de una persona.

Respondiendo Diagnóstico

¿Cuál de las siguientes opciones es igual a $\log 12$?

- a. $\log 6 \cdot \log 2$
- b. $\log 10 + \log 2$
- c. $2 \cdot \log 6$
- d. $\log 2 \cdot \log 2 \cdot \log 3$
- e. $\log 6 + \log 2$

Responder

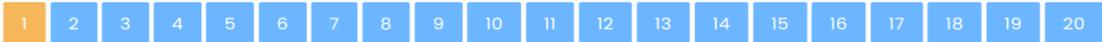


Figura 6: Ejemplo de Vista de pregunta del diagnóstico - Elaboración Propia

4.8.2 Vista de Entrenamiento

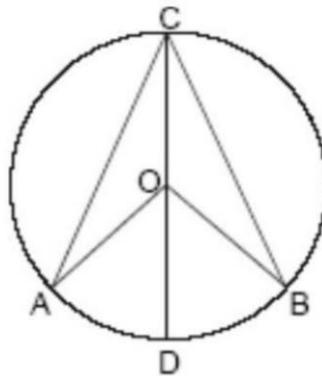
En la figura 7 se puede ver la vista de entrenamiento. Allí se muestra el número de la pregunta para poder tener un seguimiento en caso de haber algún problema. También muestra la pregunta con sus respectivas alternativas. Finalmente, en la parte inferior se tiene el botón responder. Es importante notar, que, como se había mencionado en la sección de "descripción del algoritmo" la alternativa que se marca se guarda como evidencia, independiente de si es la que se responde definitivamente o no.

Entrenamiento

Pregunta 822

En la figura, \overline{CD} es un diámetro de la circunferencia de centro O . Si el $\angle BOD = 20^\circ$ y arco AD es congruente con el arco DB , entonces ¿cuál(es) de las siguientes afirmaciones es (son) **FALSA(S)**?

- i. $\angle CBO = 20^\circ$
- ii. $\angle CAO = \angle AOD$
- iii. $\angle AOD = \angle BOD$



- a. Sólo I
- b. Sólo II
- c. Sólo I y II
- d. Sólo II y III
- e. I, II y III

Responder

Figura 7: Ejemplo de Vista de pregunta del entrenamiento - Elaboración Propia

4.9 Tecnologías Utilizadas

Como servidor web se utilizó Apache¹³. Para la aplicación web se utilizó principalmente PHP¹⁴ con Phalcon¹⁵, un framework MVC. En cuanto al motor

¹³ Apache HTTP Server Version 2.4 Documentation - Apache HTTP Server Version 2.4
<https://httpd.apache.org/docs/2.4/>

¹⁴ PHP: PHP 7.3.0 Release Announcement https://www.php.net/releases/7_3_0.php

¹⁵ Phalcon Documentation <https://docs.phalcon.io/3.4/en/introduction>

recomendador, se programó en Kotlin¹⁶, utilizando Spring Boot¹⁷ para generar el web service. Finalmente, para la base de datos se utilizó MariaDB¹⁸.

Potencialmente se podría hacer todo desde el mismo PHP, incorporando los algoritmos en el mismo lugar, pero se necesita que la parte de procesamiento, que es lo más costoso en términos de procesamiento, sea rápido. La idea es que en cuanto una persona mande una pregunta y se le envíe una de vuelta, se haga el procesamiento en tiempo real de la pregunta que sigue. Entonces para lograr que ese rendimiento sea mejor, la idea no es programar el recomendador en PHP, sino que, programarlo en Kotlin, porque Kotlin es un lenguaje que es un poco más de bajo nivel que PHP, por lo que es más eficiente.

Respecto a utilizar un lenguaje de más bajo nivel, es importante hablar sobre Kotlin y Java previamente. Ambos son lenguajes que se compilan y se ejecutan en la Java Virtual machine. Eso quiere decir que se podría ocupar cualquiera de los dos, ya que los dos están funcionando en la Java Virtual Machine una vez que son compilados, pero Kotlin es más atractivo ya que es más sencillo de utilizar, acercándolo a lenguajes como Python. Por ejemplo, un código hecho en Kotlin es más fácil de entender que uno hecho en Java, entonces se tiene una programación más cercana al humano por ser más simple su código, pero también es más eficiente ya que como es un lenguaje compilado está más cerca de la máquina.

¹⁶ Kotlin 1.3.50 released | The Kotlin Blog <https://blog.jetbrains.com/kotlin/2019/08/kotlin-1-3-50-released/>

¹⁷ <https://spring.io/blog/2018/07/30/spring-boot-2-0-4-available-now>

¹⁸ MariaDB 10.4.7 Release Notes - MariaDB Knowledge Base <https://mariadb.com/kb/en/mariadb-1047-release-notes/>

El siguiente diagrama ayuda a comprender de una manera más completa el funcionamiento de todo el sistema, incluyendo la aplicación web, el web service, y la base de datos.

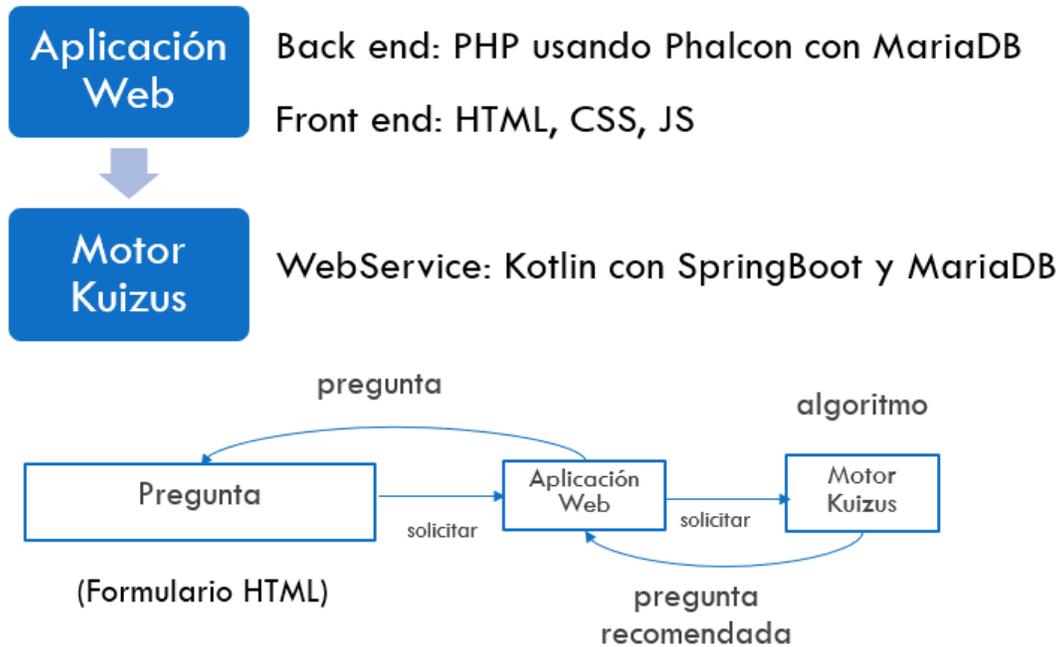


Figura 8: Arquitectura Tecnológica - Elaboración Propia

Capítulo 5: Evaluación de Sistema Kuizus

Para verificar si es que el sistema de recomendación PAES puede mejorar más los resultados de los estudiantes que un sistema de ejercitación tradicional, se realizará una evaluación.

5.1 Diseño de Evaluación

La evaluación considerará dos grupos. Un grupo experimental, y un grupo de control. Idealmente ambos grupos deben ser de cuartos medios. El grupo experimental se enfrentará a las preguntas con las recomendaciones del sistema, y el grupo de control recibirá preguntas, pero sin que incluyan los algoritmos de recomendación, serán simplemente preguntas que se van generando de manera aleatoria, de manera de poder comparar el desempeño de quienes estuvieron sometidos a las recomendaciones, versus quienes ejercitaron de manera más tradicional (sin elegir las preguntas de antemano).

Antes de comenzar a utilizar el sistema de recomendación, se les aplicará una prueba de diagnóstico a ambos grupos, tipo ensayo PAES, esto quiere decir que las preguntas tratarán de abordar todos los contenidos, pero en un formato más reducido, 20 preguntas. Que como se vio en la sección de sistemas de recomendación, lo ideal, es que al menos se enfrenten a la clasificación de 20 items (en este caso a resolver 20 preguntas) para poder empezar a recibir recomendaciones. Esta será la información que se utilizará como resultado inicial para la comparación al final de la evaluación. Luego, 4 semanas más tarde, se les volverá a tomar una prueba de diagnóstico a ambos grupos, información que se utilizará como el resultado final para la comparación y conclusión de la evaluación del sistema.

A ambos grupos se les dirá que hay un diagnóstico inicial y final, y que lo que se busca es ver si se mejora a lo largo de este periodo con ellos estudiando/ejercitando por su cuenta, por un periodo diario definido y una frecuencia determinada a la semana, durante el periodo que dure la

evaluación. Los estudiantes no saben si están en el grupo de control o en el grupo experimental.

5.2 Validación de la Evaluación

Para poder validar que la forma en que se estructuran las preguntas para mostrar que hubo un aumento es correcta, se recurrió a dos investigaciones, [31] y [32]. En ellas se explica por qué las preguntas de opción múltiple son una buena alternativa para medir el aprendizaje estudiantil, y también cómo deben ser estructuradas las preguntas para poder evaluar de manera efectiva el aprendizaje de los estudiantes.

Hay características que las preguntas de alternativas deben cumplir en su construcción para poder cumplir con el objetivo planteado. Su descripción [31] a continuación:

- No deben centrarse en memorizar hechos, sino que requerir de razonamiento por parte del estudiante.
- El enunciado de cada pregunta debe ser breve, y preciso, esto implica no incluir información que no sea relevante para contestar la pregunta. Una pregunta que está bien planteada se puede responder sin necesidad de ver las alternativas.
- Las opciones que se brindan deben ser del mismo tipo tanto en su forma como en el contenido que entregan. Deben ser respuestas posibles, que no sean ilógicas ni que se contradigan con las otras alternativas presentadas.
- Todas las preguntas deben tener cinco alternativas.
- Dentro de las opciones que se presentan no se deben incluir alternativas como "todas las anteriores", "ninguna de las anteriores", "alternativas a y b". En general ninguna que mencione otras alternativas.
- Las preguntas deben ser redactadas de manera afirmativa, y no preguntando cuál no es la correcta, o algo que no sea de forma positiva.

Entre las principales conclusiones que se levantan de ambas investigaciones, destaca que al considerar todos los factores que se evalúan en ambas investigaciones (muestra de conocimientos y destrezas, dificultad para escribir preguntas de opciones múltiples, la inconsistencia en la calificación de las preguntas tipo ensayo, y la integridad de las calificaciones de los exámenes), el uso del formato de las preguntas de opción múltiple para medir el aprendizaje de conocimiento gana la preferencia de acuerdo a los argumentos señalados.

5.3 Metodología de la Evaluación

La evaluación fue realizada en un total de 22 personas, cursando tercero o cuarto medio, en un lapso de tres semanas. Se comenzó el lunes 7 de junio y se finalizó el martes 29 de junio y se dividió en tres partes:

- Diagnóstico Inicial: 20 preguntas realizadas el lunes 7 de junio.
- Entrenamiento: realizado entre el martes 8 de junio y lunes 28 de junio
- Diagnóstico Final: 20 preguntas realizadas el martes 29 de junio

5.3.1 Grupo Experimental y Grupo de Control

En la evaluación se utilizaron once personas para el grupo de control y once personas para el grupo experimental. A las personas que formaron parte de estos grupos no se les dijo en qué grupo estaban, y se les solicitó que al menos utilizaran el sistema 5 horas a la semana. Es por esto que no hay un número definido de preguntas para la parte de entrenamiento, ya que dado que lo que tardan las personas en responder cada pregunta es diferente para cada uno, el número final de preguntas que responden no necesariamente coincidiría, a pesar de dedicarle más o menos el mismo tiempo.

5.3.2 Parámetros a Considerar

Los parámetros que se consideraron se dividen entre los concernientes a las preguntas en sí, y a los de las respuestas entregadas por los usuarios. Se

consideraron la cantidad de preguntas realizadas en total, por cada habilidad y contenido, y el porcentaje de respuestas correctas en la iteración inicial y en la iteración final.

Capítulo 6: Resultados

6.1 Resultados Obtenidos

Aquí se detallarán los resultados de la realización de la evaluación mencionada en la sección anterior. Los resultados obtenidos de esta evaluación se dividirán en tres partes:

- Diagnóstico inicial para el grupo de control y el grupo experimental.
- Entrenamiento para ambos grupos.
- Diagnóstico final para ambos grupos.

6.1.1 Diagnóstico Inicial

Los resultados del diagnóstico inicial para el grupo de control y para el grupo experimental pueden verse en la siguiente tabla:

Usuarios Grupo Control	Respuestas Correctas	Usuarios Grupo Experimental	Respuestas Correctas
Usuario 1	10 respuestas	Usuario 12	14 respuestas
Usuario 2	14 respuestas	Usuario 13	15 respuestas
Usuario 3	13 respuestas	Usuario 14	11 respuestas
Usuario 4	18 respuestas	Usuario 15	11 respuestas
Usuario 5	11 respuestas	Usuario 16	10 respuestas
Usuario 6	10 respuestas	Usuario 17	18 respuestas
Usuario 7	14 respuestas	Usuario 18	15 respuestas
Usuario 8	12 respuestas	Usuario 19	13 respuestas
Usuario 9	8 respuestas	Usuario 20	10 respuestas
Usuario 10	11 respuestas	Usuario 21	14 respuestas
Usuario 11	13 respuestas	Usuario 22	11 respuestas

Tabla 6: Resultado Preguntas Diagnóstico Inicial - Elaboración Propia

De acuerdo a lo que puede observarse el promedio de respuestas correctas en el caso del grupo de control es de 12,18 y en el caso del grupo experimental es de 12,9.

6.1.2 Entrenamiento

Antes de que los usuarios seleccionen la respuesta definitiva que escogerán para cada pregunta, pueden seleccionar otras alternativas. El promedio de las alternativas por las cuales se pasó antes de contestar la respuesta definitiva fue de 0,72 para el grupo de control y de 0,91 para el grupo experimental.

En la tabla se muestra cuántas preguntas se realizaron de cada tema. Se agruparon de esta forma para poder mostrar los resultados más fácilmente, ya que los contenidos totales para los cuales se realizó el entrenamiento eran 43, es por eso que en la siguiente tabla se muestran los seis temas que agrupan esos 43 contenidos y cuántas fueron las preguntas realizadas por cada uno de ellos.

Contenidos	Grupo de Control	Grupo Experimental
Números	255	305
Álgebra	309	316
Funciones	236	293
Geometría Posicional y Métrica	247	281
Geometría Proporcional	320	298
Datos y Azar	290	302
Total preguntas	1657	1792

Tabla 7: Preguntas realizadas por tema en Entrenamiento - Elaboración Propia

En promedio cada usuario respondió 151 preguntas en total en el grupo de control (las preguntas aleatorias), y 163 preguntas para el caso del grupo experimental (las preguntas recomendadas por el sistema). Estas preguntas se distribuyeron de la siguiente forma en el grupo de control y en el grupo experimental:

Usuarios Grupo Control	Preguntas Realizadas	Usuarios Grupo Experimental	Preguntas Realizadas
Usuario 1	193	Usuario 12	213

Usuario 2	142	Usuario 13	161
Usuario 3	151	Usuario 14	163
Usuario 4	205	Usuario 15	148
Usuario 5	156	Usuario 16	131
Usuario 6	122	Usuario 17	179
Usuario 7	139	Usuario 18	125
Usuario 8	111	Usuario 19	301
Usuario 9	156	Usuario 20	103
Usuario 10	175	Usuario 21	126
Usuario 11	107	Usuario 22	145

Tabla 8: Resultado Preguntas Entrenamiento - Elaboración Propia

6.1.3 Diagnóstico Final

Los resultados del diagnóstico final para el grupo de control y para el grupo experimental pueden verse a continuación:

Usuarios Grupo Control	Respuestas Correctas	Usuarios Grupo Experimental	Respuestas Correctas
Usuario 1	14 respuestas	Usuario 12	16 respuestas
Usuario 2	16 respuestas	Usuario 13	17 respuestas
Usuario 3	18 respuestas	Usuario 14	16 respuestas
Usuario 4	18 respuestas	Usuario 15	14 respuestas
Usuario 5	11 respuestas	Usuario 16	12 respuestas
Usuario 6	12 respuestas	Usuario 17	17 respuestas
Usuario 7	9 respuestas	Usuario 18	19 respuestas
Usuario 8	15 respuestas	Usuario 19	16 respuestas
Usuario 9	11 respuestas	Usuario 20	8 respuestas
Usuario 10	11 respuestas	Usuario 21	14 respuestas
Usuario 11	12 respuestas	Usuario 22	14 respuestas

Tabla 9: Resultado Preguntas Diagnóstico Final - Elaboración Propia

En el caso del grupo de control, el promedio de respuestas correctas fue 13,36 y para el grupo experimental fue de 14,81.

6.2 Análisis de Resultados

De acuerdo con lo que se observa en los resultados del entrenamiento, los contenidos más recomendados en el caso del grupo experimental fueron los de Álgebra, y en el caso del grupo de control, las preguntas aleatorias que más se repitieron fueron las de Geometría Proporcional.

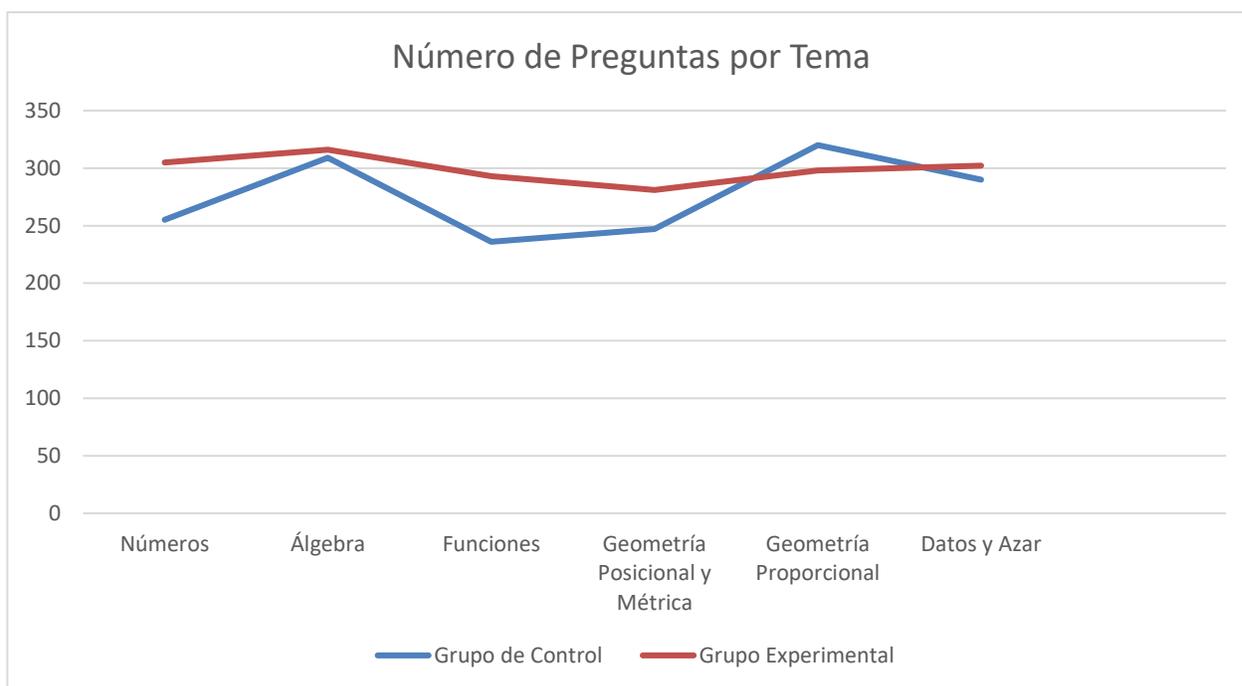


Figura 9: Número de Preguntas por Tema - Elaboración Propia

Sin embargo, puede apreciarse que la variación que hubo entre los temas es mayor en el grupo de control, donde las preguntas se realizaron de manera aleatoria, que la que hubo en el grupo experimental, donde las preguntas que tenían que responder los usuarios se fueron realizando en base a las recomendaciones realizadas por el sistema. Lo cual probablemente se debe a que en general los estudiantes tenían un conocimiento más o menos parejo de los temas, lo que termina haciendo que el número de preguntas por cada tema se normalice.

En cuanto al resultado en el cambio de los diagnósticos, a continuación, se pueden ver las variaciones para ambos grupos:

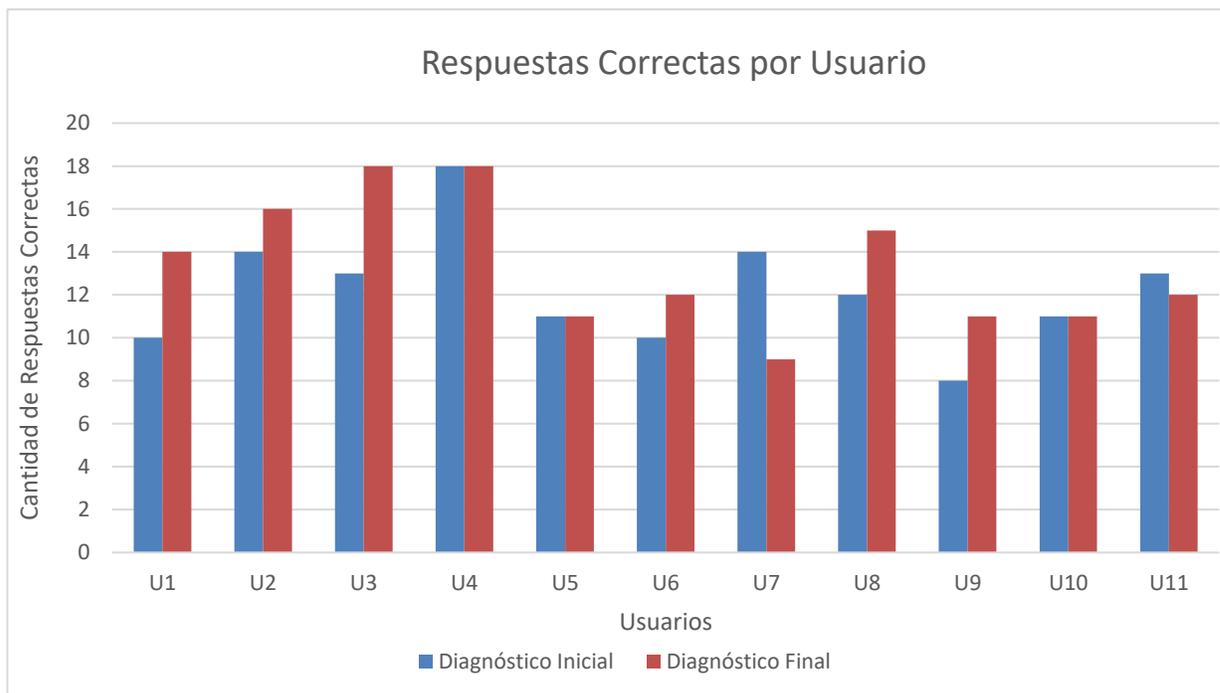


Figura 10: Respuestas correctas de los diagnósticos del grupo de control - Elaboración Propia

Para el grupo de control el promedio de las respuestas correctas pasó de 12,18 a 13,36, subiendo un 9,7%.

En el caso del grupo experimental el promedio de las respuestas correctas pasó de 12,9 a 14,81, es decir, subió un 14.8%.

La medida de comparación utilizada para observar si la diferencia entre ambos grupos es estadísticamente significativa es la de t de Student. Se utiliza esta medida, por la naturaleza de las muestras (pequeñas y asumiendo que tienen distribución normal) [33].

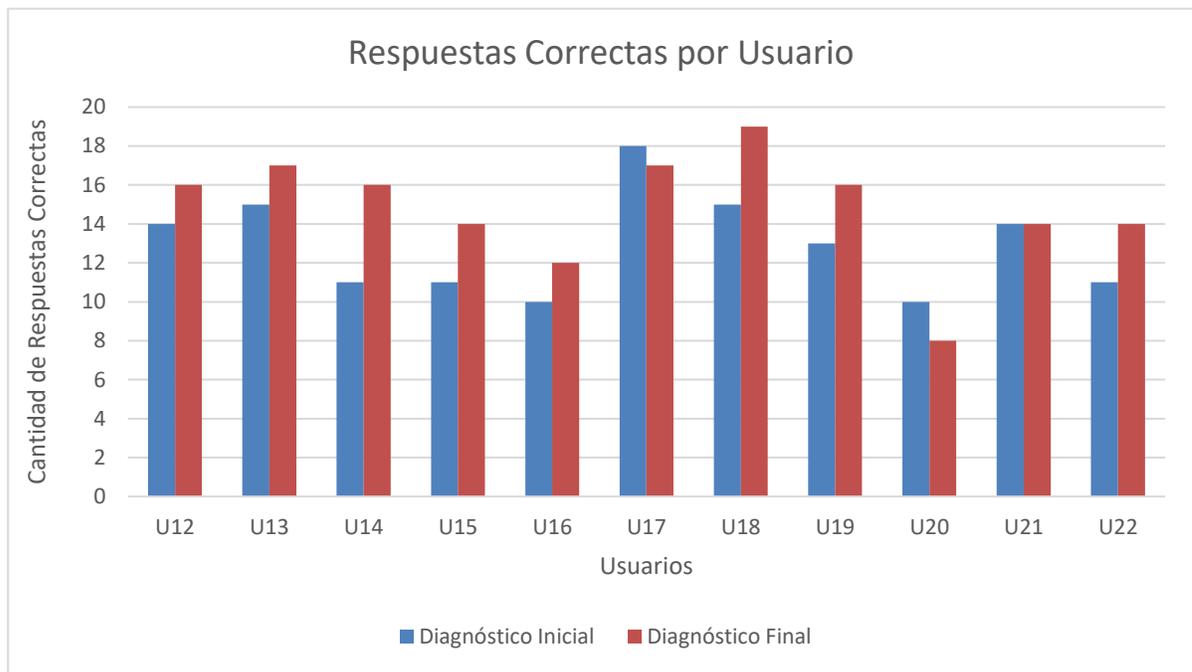


Figura 11: Respuestas correctas de los diagnósticos del grupo experimental - Elaboración Propia

Las hipótesis a contrastar con el test son las siguientes:

- Hipótesis Nula H_0 = las dos muestras no son diferentes
- Hipótesis Alternativa H_1 = las dos muestras son diferentes

Para el diagnóstico inicial los valores para calcular la t de Student fueron:

- media del grupo de control \bar{x} : 12,1818
- media del grupo experimental \bar{y} : 12,909
- desviación estándar del grupo de control σ_x : 2,6765
- desviación estándar del grupo experimental σ_y : 2,5477
- tamaño de ambas muestras $n_1 + n_2$: 22

Lo que de acuerdo a la siguiente fórmula para calcular la t de Student, entrega el valor de T a continuación:

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}}} = 1,0552$$

Para saber si la hipótesis nula se cumple, se compara este valor con el estadístico de t, que considera los siguientes parámetros:

- confianza $(1 - \alpha) = 95\%$
- grados de libertad $(n - 1) = 21$

De acuerdo a la tabla de Distribución t de Student del Anexo C, este valor es "1,721", lo que indica que ambas muestras no son significativamente diferentes ya que $1,0552 < 1,721$. Luego al realizar la misma prueba en las muestras luego de utilizar el sistema Kuizus, los valores para calcular la t de student fueron:

- media del grupo de control \bar{x} : 13,3636
- media del grupo experimental \bar{y} : 14,8182
- desviación estándar del grupo de control σ_x : 3,0421
- desviación estándar del grupo experimental σ_y : 2,9603
- tamaño de ambas muestras $n_1 + n_2$: 22

Entregando como valor:

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}}} = 1,9691$$

Como $1,9691 > 1,721$ se puede concluir que las muestras son significativamente diferentes.

Este resultado puede deberse al hecho que, dado que ambos grupos fueron sometidos a una ejercitación extra, era esperable que en ambos casos aumentaran sus resultados, por simplemente tener un estímulo que no se tenía (probablemente podría haber sido de la misma forma ante cualquier otro método de ejercitación). Sin embargo, destaca el hecho que el grupo experimental tuvo una mejora mayor que el grupo de control, testeada con el test t de Student, caracterizada como significativamente diferente, lo que se relaciona con el hecho de haber utilizado el algoritmo de recomendación, ya que las preguntas que se recibieron fueron dirigidas especialmente a los usuarios basados en sus conocimientos particulares, y no de manera aleatoria o estándar.

Capítulo 7: Conclusiones

7.1 Acerca de la Hipótesis de Investigación

Los resultados muestran una mejora significativa del grupo experimental respecto al diagnóstico inicial en comparación con el grupo de control, se puede concluir que la diferencia es significativa gracias al test t de Student realizado. Por lo que se concluye que la hipótesis de investigación “Los estudiantes pueden obtener mejores resultados en su PAES de matemáticas si utilizan como ejercitación el método Kuizus” se cumple.

Es importante mencionar que en promedio el grupo experimental respondió un promedio de 7.9% más de preguntas, lo que tal vez podría haber influenciado su resultado ya que ejercitaron un poco más que el grupo de control.

7.2 Acerca de la Metodología Utilizada en el Algoritmo Recomendador

Respecto a la metodología utilizada en el algoritmo recomendador se puede concluir que las recomendaciones que se realizaron fueron buenas, debido a la mejora que se vio en los resultados del grupo que las tuvo, versus el que no. Sin embargo, la data que se tiene actualmente en el sistema no fue suficiente para poder hacer además recomendaciones basadas en los clusters de usuarios que se fueron generando en el sistema. Es esperable que con el tiempo y por consecuencia mayor cantidad de datos acumulados, se pueda hacer mejor esta clusterización, pero de momento no se logró una agrupación que fuera lo suficientemente potente.

Por otro lado, las recomendaciones en base al historial de respuestas del mismo usuario resultaron exitosas, y se mostró mejora en el aprendizaje. Aunque en el caso del grupo que recibía respuestas en forma aleatoria también hubo mejora, en el caso de las preguntas recomendadas por el sistema fue un 52% mayor. Esto hace sentido, ya que muestra que el aprendizaje por medio de ensayo y error lleva finalmente a la adquisición de

nuevos conocimientos, debido a que se van reforzando los conocimientos más débiles.

7.3 Acerca de los Objetivos

Los objetivos de esta tesis fueron logrados, aunque, como se mencionó anteriormente, se necesitan más datos para poder probar de una forma más robusta la hipótesis de investigación (incluyendo esta vez recomendaciones basadas en clusterización entre pares), si se logró mostrar que una herramienta como esta ayuda en el aprendizaje de los alumnos. Probablemente el solo hecho de estar expuesto a este tipo de herramientas, ayuda en el estudio, lo que permite tener un mejor aprendizaje, pero como se pudo ver en el grupo de personas evaluadas, quienes fueron recomendados a repetir los contenidos donde sus conocimientos eran más débiles, pudieron lograr tener eventualmente un mayor conocimiento con el diagnóstico final.

En cuanto a los temas que trataban las preguntas, se puede ver como el tipo de contenido que tenía las preguntas empieza a converger en el tiempo. Esto muestra que las distintas temáticas se empezaron a acercar en su nivel de conocimiento por parte de los usuarios.

7.4 Trabajo Futuro

El trabajo futuro de esta tesis puede clasificarse en dos formas distintas, a nivel algorítmico y de arquitectura.

- Respecto a la materia algorítmica, la data capturada como evidencia no fue utilizada en la presente tesis, pero podría ser utilizada para refinar el motor recomendador a medida que la base de datos se va volviendo más grande. Por ejemplo, asociar una mayor dificultad a las preguntas que toman más tiempo en ser respondidas, y por ello que queden más propensas a ser recomendadas.

- En cuanto a la arquitectura del sistema presentado, dado el tamaño actual de la base de datos, se logra recomendar una pregunta en cerca de 0.7 segundos, pero eventualmente al complejizarse el modelo, podría tener un deterioro en la performance. También se debe tener en consideración que a medida que se tengan más datos podría ir ocurriendo algo similar. Dicho lo anterior, se recomienda estudiar dos aproximaciones de mejora:
 - Utilizar tablas materializadas para evitar cálculos. Esto ayudaría a que los resultados sean más rápidos, ya que se tendrían ciertas partes pre-calculadas.
 - Utilizar una estructura de procesamiento en stream distribuida. Con esto, se podrían utilizar varias máquinas para que hagan el procesamiento en paralelo de las distintas recomendaciones. Esto último no es prioridad en el estado actual de este proyecto, pero podría llegar a ser necesario si las preguntas en la tabla de respuestas crecen de manera acelerada.

Capítulo 8: Bibliografía

- [1] Acción Educar, «Alumnos por curso en la educación escolar: realidad actual y su relación con los resultados de aprendizaje,» Acción Educar, Santiago, 2015.
- [2] K. Falk, «The Netflix recommender system,» de *Practical Recommender Systems*, Shelter Island, Manning, 2019, p. 6.
- [3] K. Falk, «Taxonomy of recommender systems,» de *Practical Recommender Systems*, Shelter Island, Manning, 2019, pp. 15-22.
- [4] R. Burke, «Recommendation Techniques,» *User Modeling and User-Adapted Interaction*, vol. 12, pp. 2-3, 2002.
- [5] K. Falk, «Non-personalized recommendations,» de *Practical Recommender Systems*, Shelter Island, Manning, 2019, p. 115.
- [6] K. Falk, «The user (and content) who came in from the cold,» de *Practical Recommender Systems*, Shelter Island, Manning, 2019, p. 130.
- [7] K. Falk, «What are implicit ratings?,» de *Practical Recommender Systems*, Shelter Island, Manning, 2019, p. 85.
- [8] K. Falk, «Time-Based Approach,» de *Practical Recommender Systems*, Shelter Island, Manning, 2019, pp. 86-87.
- [9] D. G. McIlwraith, H. Marmanis y D. Babenko, «Distance and similarity,» de *Algorithms of the Intelligent Web*, Shelter Island, Manning, 2016, p. 49.
- [10] E. Kalayci, «The trade-off in machine learning; Accuracy vs explainability,» 25 diciembre 2018. [En línea]. Available: <https://medium.com/@erdemkalayci/the-tradeoff-in-machine-learning-accuracy-vs-explainability-fbb13914fde2>. [Último acceso: 1 abril 2020].
- [11] K. Falk, «Output,» de *Practical Recommender Systems*, Shelter Island, Manning, 2019, pp. 20-21.
- [12] K. Falk, «User behavior and how to collect it,» de *Practical Recommender Systems*, Shelter Island, Manning, 2019, p. 30.
- [13] K. Falk, «Finding useful user behavior,» de *Practical Recommender Systems*, Shelter Island, Manning, 2019, p. 35.
- [14] K. Falk, «Essential similarity function,» de *Practical Recommender Systems*, Shelter Island, Manning, 2019, p. 152.

- [15] K. Falk, «What's a similarity function?,» de *Practical Recommender Systems*, Shelter Island, Manning, 2019, p. 153.
- [16] K. Falk, «Table Different data types,» de *Practical Recommender Systems*, Shelter Island, Manning, 2019, p. 154.
- [17] K. Falk, «Jaccars distance,» de *Practical Recommender Systems*, Shelter Island, Manning, 2019, pp. 155-156.
- [18] K. Falk, «L1 Norm,» de *Practical Recommender Systems*, Shelter Island, Manning, 2019, pp. 157-158.
- [19] K. Falk, «Cosine similarity,» de *Practical Recommender Systems*, Shelter Island, Manning, 2019, p. 159.
- [20] K. Falk, «Finding similarity with Pearson's correlation coefficient,» de *Practical Recommender Systems*, Shelter Island, Manning, 2019, p. 162.
- [21] C. Krauss, A. Merceron y S. Arbanowski, «Measurement Values,» de *The Timeliness Deviation: A novel Approach to Evaluate Educational Recommender Systems for Closed-Courses*, New York, 2019.
- [22] F. Anaraki, «Developing an Effective and Efficient eLearning Platform,» *International Journal of The Computer, the Internet and Management*, vol. 12, pp. 57-63, 2004.
- [23] S. Burd, *Web-based support of a programming class*, Hershey: Idea Group Publishing.
- [24] O. R. Zaïane, «Building a Recommender Agent for e-Learning Systems,» de *International Conference on Computers in Education*, Alberta, 2002.
- [25] S. Shee, J. Han y K. Wang, «An efficient collaborative filtering method,» de *Conference on Data Warehousing and Knowledge Discovery*, Germany, 2001.
- [26] O. R. Zaïane, *Web usage mining for a better web-based learning environment.*, Banff: Conference on Advanced Technology for Education, 2001.
- [27] S. AI, «Squirrel AI Learning,» Squirrel AI Learning, [En línea]. Available: <http://squirrelai.com>. [Último acceso: 12 05 2020].
- [28] K. Hao, «China has started a grand experiment in AI education,» MIT Technology Review, [En línea]. Available: <https://www.technologyreview.com/2019/08/02/131198/china-squirrel-has-started-a-grand-experiment-in-ai-education-it-could-reshape-how-the/>. [Último acceso: 12 05 2020].
- [29] A. M. Antelm Lanzat, A. J. Gil Lopez, M. L. Cacheiro González y E. Pérez Navío, *Causas del fracaso escolar: un análisis desde la perspectiva del profesorado y del alumnado*, Enseñanza & Teaching Revista Interuniversitaria de Didáctica, 2018.

- [30] DEMRE, *Temario PAES regular de Competencia Matemática 1 (M1)*, Santiago: DEMRE, 2023.
- [31] M. T. Chiang y C. Díaz, «Generalidades de Evaluación y Elaboración de Preguntas de Opción Múltiple,» Dirección de Docencia - Universidad de Concepción, Concepción, 2011.
- [32] Educrea, «Preparación de preguntas de opciones múltiples para medir el aprendizaje de los estudiantes,» Educrea, [En línea]. Available: <https://educrea.cl/preparacion-de-preguntas-de-opciones-multiples-para-medir-el-aprendizaje-de-los-estudiantes/>. [Último acceso: 1 Mayo 2020].
- [33] B. Fadem, *High-Yield Behavioral Science*, Hagerstown: Lippincott Williams & Wilkins, 2008.
- [34] E. Fasce y P. Ibáñez, *Apunte Taller Confeción de Preguntas de Opción Múltiple*, Concepción: Universidad de Concepción, 2007.

Anexos

Anexo A: Temario PAES Matemáticas

CONOCIMIENTOS EVALUADOS EN LA PRUEBA DE MATEMÁTICA 1

EJE TEMÁTICO: NÚMEROS

- Conjunto de los números enteros y racionales:
 - Operaciones y orden en el conjunto de los números enteros y racionales.
 - Problemas que involucren el conjunto de los números enteros y racionales en diversos contextos.
- Porcentaje:
 - Concepto y cálculo de porcentaje.
 - Problemas que involucren porcentaje en diversos contextos.
- Potencias y raíces enésimas:
 - Propiedades de las potencias de base racional y exponente racional.
 - Descomposición y propiedades de las raíces enésimas en los números reales.
 - Problemas que involucren potencias y raíces enésimas en los números reales en diversos contextos.

EJE TEMÁTICO: ALGEBRA Y FUNCIONES

- Expresiones algebraicas:
 - Productos notables.
 - Factorizaciones de expresiones algebraicas.
 - Operatoria con expresiones algebraicas.
 - Problemas que involucren expresiones algebraicas en diversos contextos.
- Proporcionalidad:

- Concepto de proporción directa e inversa con sus diferentes representaciones.
- Problemas que involucren proporción directa e inversa en diversos contextos.
- Ecuaciones e inecuaciones de primer grado:
 - Resolución de ecuaciones lineales.
 - Problemas que involucren ecuaciones lineales en diversos contextos.
 - Resolución de inecuaciones lineales.
 - Problemas que involucren inecuaciones lineales en diversos contextos.
- Sistemas de ecuaciones lineales (2x2):
 - Resolución de sistemas de ecuaciones lineales.
 - Problema que involucren sistemas de ecuaciones lineales en diversos contextos.
- Función lineal y afín:
 - Concepto de función lineal y función afín.
 - Tablas y gráficos de función lineal y función afín.
 - Problemas que involucren función lineal y función afín en diversos contextos.
- Función cuadrática:
 - Resolución y problemas de ecuaciones de segundo grado en diversos contextos.
 - Tablas y gráficos de la función cuadrática, considerando la variación de sus parámetros.
 - Puntos especiales de la gráfica de la función cuadrática: vértice, ceros de la función e intersección con los ejes.
 - Problemas que involucren la función cuadrática en diversos contextos.

EJE TEMÁTICO: GEOMETRÍA

- Figuras Geométricas:
 - Problemas que involucren el Teorema de Pitágoras en diversos contextos.

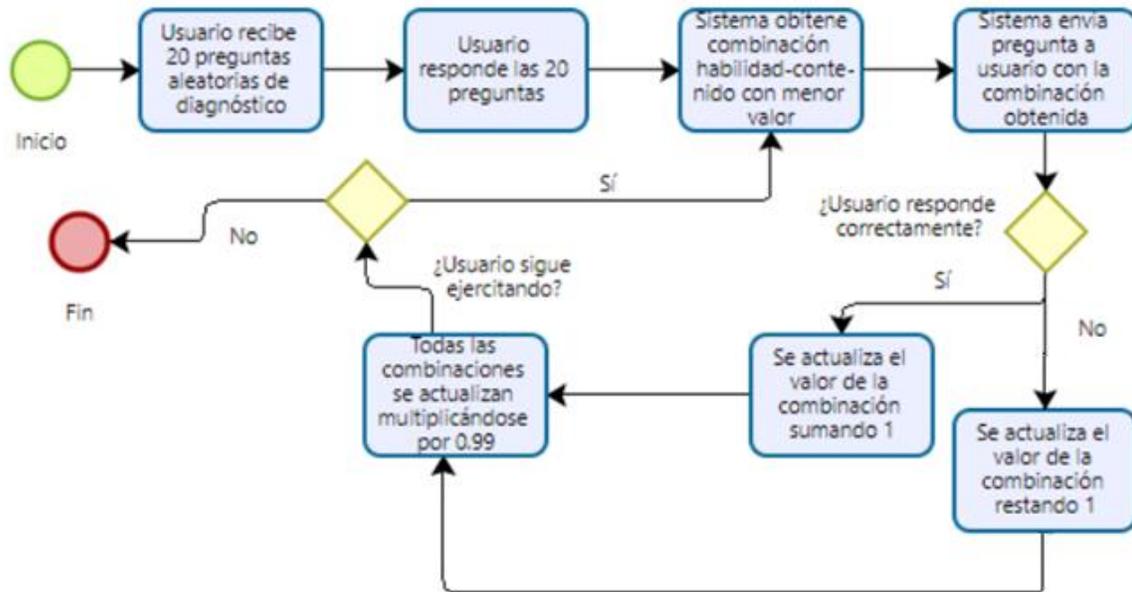
- Perímetro y áreas de triángulos, paralelogramos, trapecios, círculos, segmentos y sectores circulares.
- Problemas que involucren perímetro y áreas de triángulos, paralelogramos, trapecios, círculos, segmentos y sectores circulares en diversos contextos.
- Cuerpos geométricos:
 - Área de superficies de prismas rectos con diferentes bases y cilindros.
 - Volumen de prismas rectos con diferentes bases y cilindros.
 - Problemas que involucren área y volumen de prismas rectos y cilindros en diversos contextos.
- Transformaciones isométricas:
 - Puntos y vectores en el plano cartesiano.
 - Rotación, traslación y reflexión de figuras geométricas.
 - Problemas que involucren rotación, traslación y reflexión en diversos contextos.

EJE TEMÁTICO: PROBABILIDAD Y ESTADÍSTICA

- Representación de datos a través de tablas y gráficos:
 - Tablas de frecuencia absoluta y relativa
 - Tipos de gráficos que permitan representar datos
 - Problemas que involucren tablas y gráficos en diversos contextos.
- Medidas de tendencia central y rango:
 - Medidas de tendencia central y rango de uno o más grupos de datos.
 - Problemas que involucren medidas de tendencia central y rango en diversos contextos.
- Medidas de posición:
 - Cuartiles y percentiles de uno o más grupos de datos
 - Diagrama de cajón para representar distribución de datos.
 - Problemas que involucren medidas de posición en diversos contextos.
- Reglas de las probabilidades:

- Problemas que involucren probabilidad de un evento en diversos contextos
- Problemas que involucren la regla aditiva y multiplicativa de probabilidades en diversos contextos.

Anexo B: Algoritmo de Sistema Recomendador



Anexo C: Extracto Distribución t de Student

		$\alpha/2$												
		0,0005	0,001	0,005	0,01	0,025	0,05	0,1	0,2	0,25	0,3	0,4	0,45	0,475
v grados de libertad	1	636,619	318,309	63,657	31,821	12,706	6,314	3,078	1,376	1,000	0,727	0,325	0,158	0,079
	2	31,599	22,327	9,925	6,965	4,303	2,920	1,886	1,061	0,816	0,617	0,289	0,142	0,071
	3	12,924	10,215	5,841	4,541	3,182	2,353	1,638	0,978	0,765	0,584	0,277	0,137	0,068
	4	8,610	7,173	4,604	3,747	2,776	2,132	1,533	0,941	0,741	0,569	0,271	0,134	0,067
	5	6,869	5,893	4,032	3,365	2,571	2,015	1,476	0,920	0,727	0,559	0,267	0,132	0,066
	6	5,959	5,208	3,707	3,143	2,447	1,943	1,440	0,906	0,718	0,553	0,265	0,131	0,065
	7	5,408	4,785	3,499	2,998	2,365	1,895	1,415	0,896	0,711	0,549	0,263	0,130	0,065
	8	5,041	4,501	3,355	2,896	2,306	1,860	1,397	0,889	0,706	0,546	0,262	0,130	0,065
	9	4,781	4,297	3,250	2,821	2,262	1,833	1,383	0,883	0,703	0,543	0,261	0,129	0,064
	10	4,587	4,144	3,169	2,764	2,228	1,812	1,372	0,879	0,700	0,542	0,260	0,129	0,064
	11	4,437	4,025	3,106	2,718	2,201	1,796	1,363	0,876	0,697	0,540	0,260	0,129	0,064
	12	4,318	3,930	3,055	2,681	2,179	1,782	1,356	0,873	0,695	0,539	0,259	0,128	0,064
	13	4,221	3,852	3,012	2,650	2,160	1,771	1,350	0,870	0,694	0,538	0,259	0,128	0,064
	14	4,140	3,787	2,977	2,624	2,145	1,761	1,345	0,868	0,692	0,537	0,258	0,128	0,064
	15	4,073	3,733	2,947	2,602	2,131	1,753	1,341	0,866	0,691	0,536	0,258	0,128	0,064
	16	4,015	3,686	2,921	2,583	2,120	1,746	1,337	0,865	0,690	0,535	0,258	0,128	0,064
	17	3,965	3,646	2,898	2,567	2,110	1,740	1,333	0,863	0,689	0,534	0,257	0,128	0,064
	18	3,922	3,610	2,878	2,552	2,101	1,734	1,330	0,862	0,688	0,534	0,257	0,127	0,064
	19	3,883	3,579	2,861	2,539	2,093	1,729	1,328	0,861	0,688	0,533	0,257	0,127	0,064
	20	3,850	3,552	2,845	2,528	2,086	1,725	1,325	0,860	0,687	0,533	0,257	0,127	0,063
	21	3,819	3,527	2,831	2,518	2,080	1,721	1,323	0,859	0,686	0,532	0,257	0,127	0,063
	22	3,792	3,505	2,819	2,508	2,074	1,717	1,321	0,858	0,686	0,532	0,256	0,127	0,063
	23	3,768	3,485	2,807	2,500	2,069	1,714	1,319	0,858	0,685	0,532	0,256	0,127	0,063