**INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE OCCIDENTE**

**Departamento de Procesos Tecnológicos e Industriales**

**Sustentabilidad y tecnología**

**PROYECTO DE APLICACIÓN PROFESIONAL (PAP)**

**Programa de Apoyo a la investigación I**



**4G03 Programa de Apoyo a la investigación I**

**Investigación de formas de procesar palabras y optimización del código generado para el mejoramiento de modelos alimenticios generados por (IA).**

**PRESENTAN**

Programas educativos y Estudiantes

Ing. Y Ciencia de Datos Kaleb Azael Avila Padilla

Profesor PAP: Dra. Raquel Zúñiga Rojas

Tlaquepaque, Jalisco, julio de 2024

**ÍNDICE**

Contenido

[REPORTE PAP 2](#_heading=h.gjdgxs)

[Presentación Institucional de los Proyectos de Aplicación Profesional 2](#_heading=h.30j0zll)

Resumen 3

[1.](#_heading=h.3znysh7) Ciclo participativo del Proyecto de Aplicación Profesional 3

[1.1 Entendimiento del ámbito y del contexto 4](#_heading=h.2et92p0)

[1.2 Caracterización de la organización 5](#_heading=h.tyjcwt)

[1.3 Identificación de la(s) problemática(s) 5](#_heading=h.3dy6vkm)

1.4 Planeación de alternativa(s) 10

[1.5. Desarrollo de la propuesta de mejora](#_heading=h.4d34og8) 13

[1.6. Valoración de productos, resultados e impactos](#_heading=h.2s8eyo1) 17

[1.7. Bibliografía y otros recursos](#_heading=h.17dp8vu) 18

[1.8. Anexos generales](#_heading=h.3rdcrjn) 19

[2. Productos](#_heading=h.26in1rg) 20

[3. Reflexión crítica y ética de la experiencia 7](#_heading=h.lnxbz9)

[3.1 Sensibilización ante las realidades 8](#_heading=h.35nkun2)

[3.2 Aprendizajes logrados 8](#_heading=h.1ksv4uv)

# REPORTE PAP

## Presentación Institucional de los Proyectos de Aplicación Profesional

*Los Proyectos de Aplicación Profesional (PAP) son experiencias socio-profesionales de los alumnos que desde el currículo de su formación universitaria- enfrentan retos, resuelven problemas o innovan una necesidad sociotécnica del entorno, en vinculación (colaboración) (co-participación) con grupos, instituciones, organizaciones o comunidades, en escenarios reales donde comparten saberes.*

*El PAP, como espacio curricular de formación vinculada, ha logrado integrar el Servicio Social (acorde con las Orientaciones Fundamentales del ITESO), los requisitos de dar cuenta de los saberes y del saber aplicar los mismos al culminar la formación profesional (Opción Terminal), mediante la realización de proyectos profesionales de cara a las necesidades y retos del entorno (Aplicación Profesional).*

*El PAP es un proceso acotado en el tiempo en que los estudiantes, los beneficiarios externos y los profesores se asocian colaborativamente y en red, en un proyecto, e incursionan en un mundo social, como actores que enfrentan verdaderos problemas y desafíos traducibles en demandas pertinentes y socialmente relevantes. Frente a éstas transfieren experiencia de sus saberes profesionales y demuestran que saben hacer, innovar, co-crear o transformar en distintos campos sociales.*

*El PAP trata de sembrar en los estudiantes una disposición permanente de encargarse de la realidad con una actitud comprometida y ética frente a las disimetrías sociales. En otras palabras, se trata del reto de “saber y aprender a transformar”.*

*El Reporte PAP consta de tres componentes:*

*El primer componente refiere al ciclo participativo del PAP, en donde se documentan las diferentes fases del proyecto y las actividades que tuvieron lugar durante el desarrollo de este y la valoración de las incidencias en el entorno.*

*El segundo componente presenta los productos elaborados de acuerdo con su tipología.*

*El tercer componente es la reflexión crítica y ética de la experiencia, el reconocimiento de las competencias y los aprendizajes profesionales que el estudiante desarrolló en el transcurso de su labor.*

## Resumen

## Ciclo participativo del Proyecto de Aplicación Profesional

El PAP es una experiencia de aprendizaje y de contribución social integrada por estudiantes, profesores, actores sociales y responsables de las organizaciones, que de manera colaborativa construir sus conocimientos para dar respuestas a problemáticas de un contexto específico y en un tiempo delimitado. Por tanto, la experiencia PAP supone un proceso en lógica de proyecto, así como de un estilo de trabajo participativo y recíproco entre los involucrados.

El proyecto se está desarrollando en el centro de ITESO con el departamento de DESI, MAF y DPTI en específico para completar el desarrollo de un modelo de inteligencia artificial que pueda generar nuevos productos alimenticios a base de la información que se le fue dada al modelo. El proyecto consta de dos partes que son el modelo y el poner a prueba lo que generó el modelo, es decir la generación del alimento en el laboratorio con los ingredientes proporcionados por el modelo.

La característica del el proyecto que se está realizando es experimental, el modelo está finalizado, pero se necesita el saber cómo es que funciona el modelo ya de ahí ver si este puede tener algún tipo de mejora, ahí es donde entra lo experimental, ya que se van a estar haciendo interacciones con el código o de la forma de manejo de los datos del modelo.

Objetivo general, desarrollar nuevos modelos para la generación de nuevos alimentos con base en el modelo que se desarrolló en Primavera 2024, modificando ya sea la generación del modelo o experimentando con nuevas formas de hacerlo.

## Entendimiento del ámbito y del contexto

El proyecto al abordar un tema tan delicado como lo es el alimento puede ser de gran ayuda para la ZMG y la República Mexicana, ya que en el estado de Jalisco y zona metropolitana hay mucha demanda en los alimentos si no que puede que los alimentos a los que la gente puede recurrir son de baja calidad o sin los nutrientes que se requieren, para esto el modelo puede ser una gran alternativa ya que nos puede dar alimentos con base a sus nutrientes y sabor, todo esto con las características similares al alimento que se quiere replicar.

Un ejemplo de esto son los alimentos que replican a los alimentos a base animal, como el Not Meat (carne hecha con hierbas entre otras cosas) que es un gran ejemplo de cómo funcionaria el modelo, como en su parte de recetas [1], por lo que además de poder beneficiar a la población de escasos recursos también esto puede beneficiar a la población que no consume producto animal.

El proyecto comenzó con la búsqueda de las bases de datos, para esto se definieron las que podrían ser las mejores bases de datos abiertas y de fácil acceso, esto ya que hay bases de datos que no son públicas o están bloqueadas para el uso, después de eso se empezó con el análisis de las bases de datos y la creación de las variables para el modelo utilizando las bases de datos que se seleccionaron, se siguió con el procesamiento de la información y las variables, se utilizó un modelo de clasificación de bosque aleatorio para la generación del modelo, se usó este modelo para poder clasificar los alimentos al observar su similitud, por ejemplo si tienes dos alimentos y ver qué tan similares son con el alimento a generar *Y* ,por lo que el modelo observa a los alimentos dados y si estos son similares al producto a generar da un resultado de 1 diciendo que el alimento es similar, después de eso se desarrolló el modelo y se sometió a varias pruebas, el modelo se pone a prueba con información que no se le fue dada durante su entrenamiento esto para ver si funciona de manera correcta, por ejemplo si se entrenó con un alimento y al ponerlo a prueba se le da y el modelo tiene que decir si esta es similar o no al producto a generar *Y.*

El modelo cuando ya está completo se utiliza un algoritmo que usa el modelo para que te genere los ingredientes para el alimento similar que se quiere generar, ya se si le pide que genere un producto *Y* o *Z* este te dará los ingredientes que sean más similares para la creación de estos productos.

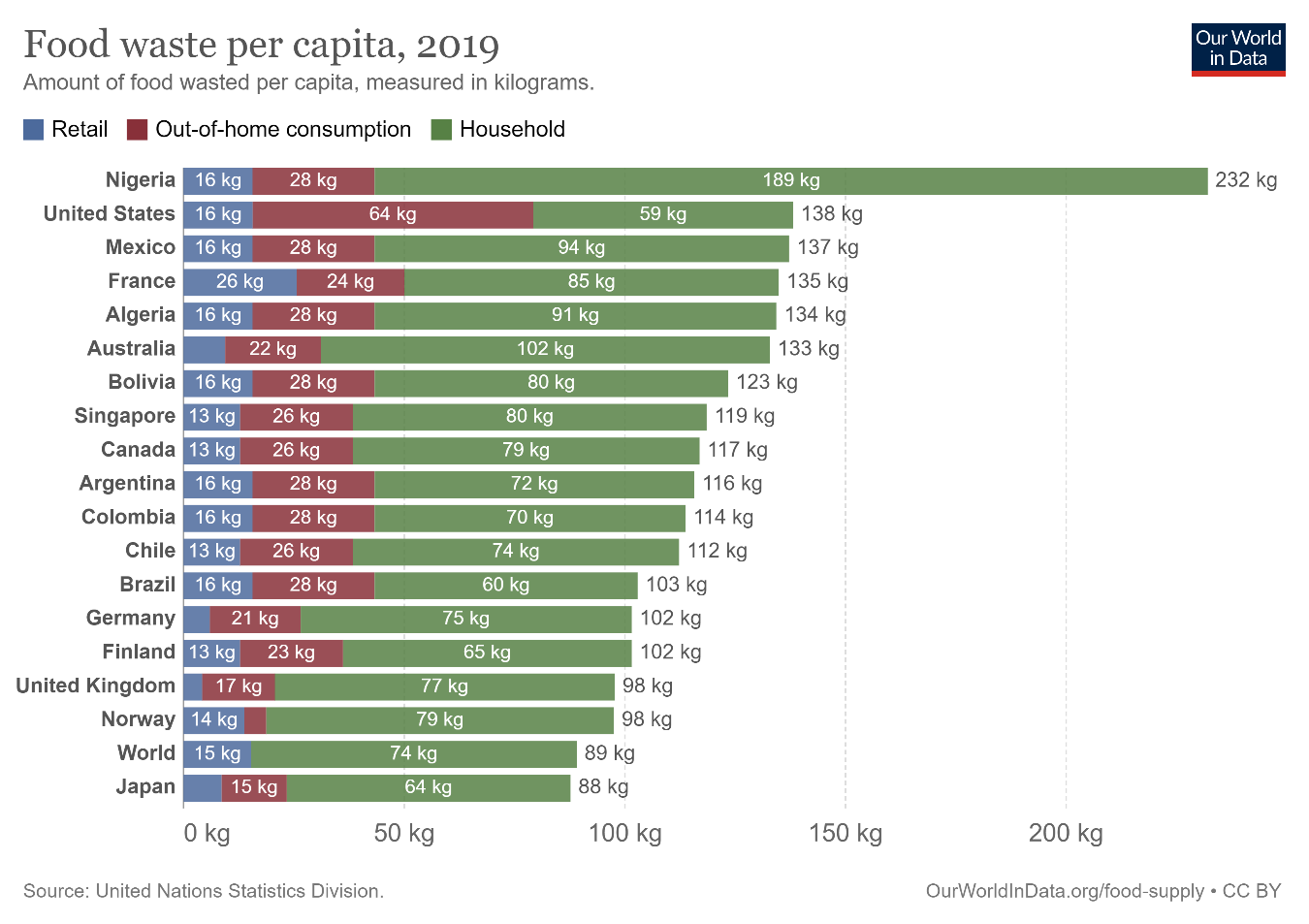
## Caracterización de la organización

El proyecto se lleva a cabo dentro del ITESO ya sea dentro del laboratorio de ingeniería de alimentos para mis compañeros o dentro de los laboratorios de computación en el T o el edificio de J donde se encuentra el departamento de MAF. En el laboratorio de ingeniería en alimentos se desarrollan los nuevos alimentos que generó el modelo, en las otras áreas más que nada se usan para investigar o mejorar el algoritmo que se usó para el modelo.

En el área de ciencia de datos hay cinco personas, Hernández Mota, Daniel es el encargado de la generación del algoritmo del modelo así como de su implementación, aunque todos estuvieron haciendo su propio modelo el de Daniel es el que se está usando ya en la fase final del proyecto, Muñoz López, Patricia es la encargada de también generar un modelo nuevo pero utilizando un nuevo enfoque, al hacer uso de las variables categóricas de las bases de datos y procesándolas con un modelo llamado Word2Vec que transforma las palabras en vectores para un mejor procesamiento de estas [2], el profesor Lozano Diaz, Cesar es el supervisor de todo los modelos generados ya que es la persona con más experiencia de todos en el uso de inteligencia artificial y modelos matemáticos, Sánchez Carrillo, Felipe se encarga de evaluar los modelos y por último Avila Padilla, Kaleb se encarga de la ayuda a todos los involucrados en los modelos, ya sea a limpiar el código, experimentar nuevas formas de abordar los datos o el generar un nuevo modelo con base a los que se han hecho.

## 1.3 Identificación de la(s) problemática(s)

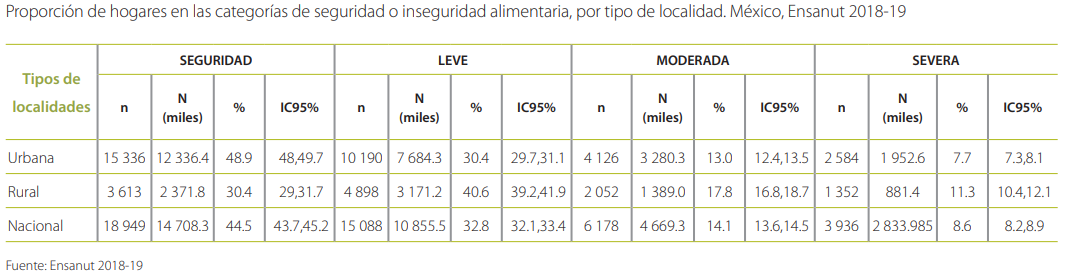
En México se desperdician alrededor de 28 millones de toneladas de alimentos al año; para 2019 hubo un desperdicio per cápita de, 16 kg de alimentos de venta al por menor, 28 kg de alimentos por consumo fuera del hogar y de 94 kg de alimentos consumidos en el hogar (Figura 1). Entre los países comparados en la imagen 2, México ocupa el 3er lugar solo después de Nigeria y Estados Unidos. Cabe mencionar que el mayor desperdicio se presenta en los alimentos consumidos en el hogar, quedando también en 3er lugar después de Nigeria y Australia. Mientras que en México se desperdiciaron en 2019 94 kg de alimentos consumidos en el hogar por habitante, en Argentina, Brasil, Chile y Colombia, este valor fue igual o estuvo por debajo del promedio mundial (74 kg/persona). De acuerdo con la Secretaría de Desarrollo Social (SEDESOL), la cantidad de alimento desperdiciado en México serviría para alimentar a más de 7 millones de mexicanos y disminuir la inseguridad alimentaria [3].



**Figura 1. Desperdicio de alimentos per cápita, medido en kilogramos: venta al por menor, consumo fuera del hogar y consumo en el hogar [3].**

Un indicador de pobreza en México desde hace más de una década ha sido la inseguridad alimentaria dentro de los hogares pues permite identificar a los grupos de la población que presentan carencia en el cumplimiento de sus derechos sociales, entre ellos la alimentación. De acuerdo con la Encuesta Nacional de Salud y Nutrición (ENSANUT) 2018-2019, en México la prevalencia de inseguridad alimentaria (Figura 2) fue de 55.5% en total, presentado la población un porcentaje de 32.8% en inseguridad leve, 14.1% en inseguridad moderada y 8.6% en inseguridad severa, es decir 18 352 241 hogares con algún grado de inseguridad alimentaria, de los cuales una mayor proporción de hogares se encontraron en localidades rurales (69.7%: 40.6% en inseguridad leve, 17.8% en moderada y 11.3% en severa) con respecto a los hogares urbanos (51.1%: 30.4% en inseguridad leve, 13.0% en moderada y 7.7% en severa) [4].

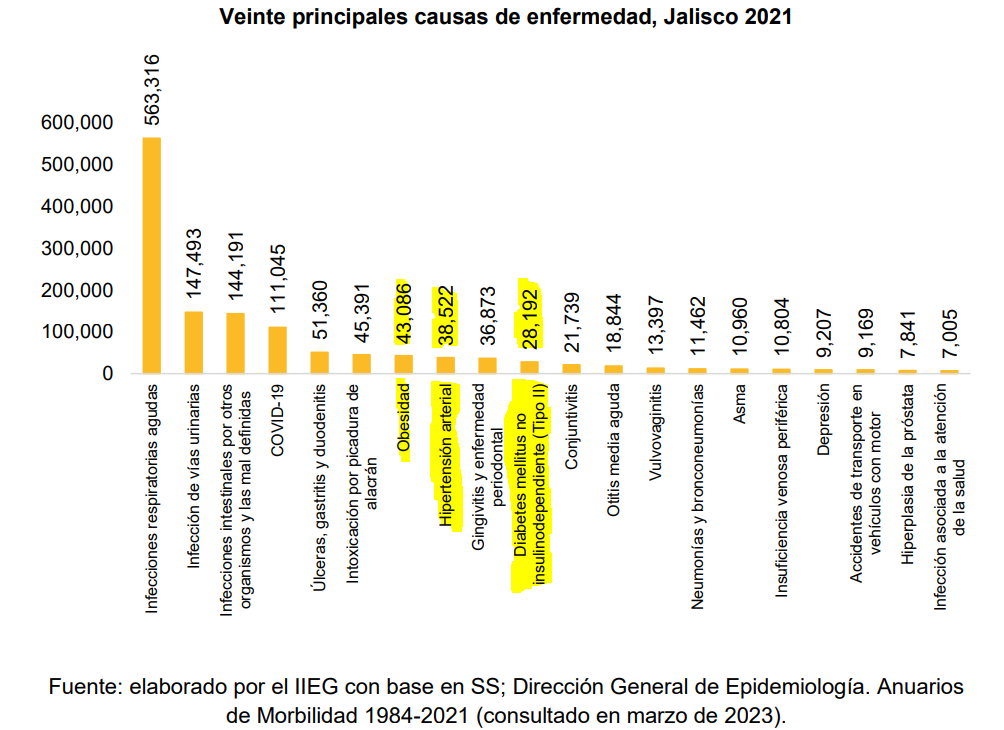
La mayoría de los hogares en inseguridad alimentaria se ubicó en las regiones Sur (68.4%) y Centro (52.8%) del país. Asimismo, la proporción más grande de hogares en las formas más graves de inseguridad alimentaria (moderada y severa) se encuentran en dichas regiones del país (29.3% en la región Sur y 21.5% en la región Centro) [4].



**Figura 2. Proporción de hogares en las categorías de seguridad o inseguridad alimentaria, por tipo de localidad en México [4].**

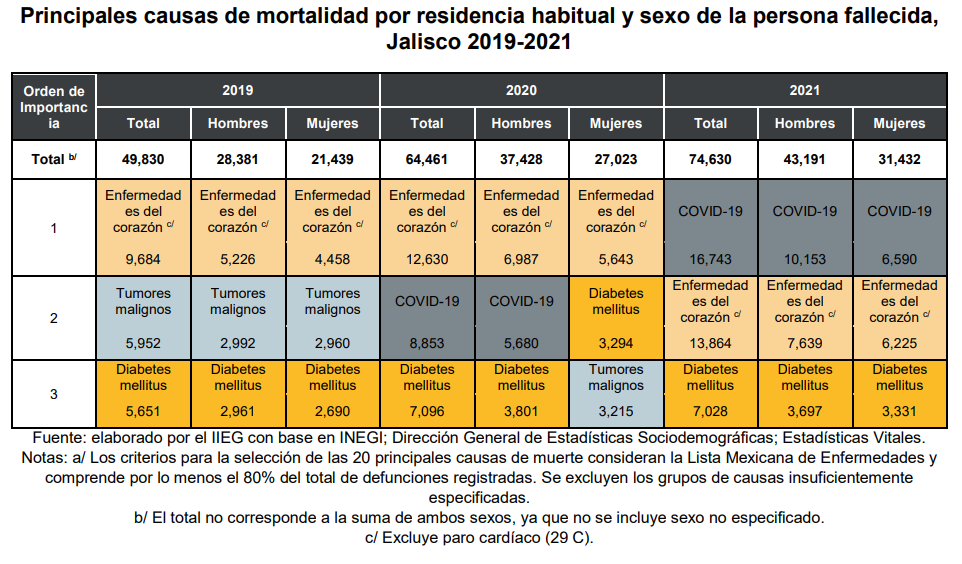
La inseguridad alimentaria, no solo está relacionada con un acceso limitado a los alimentos, sino también con un poder adquisitivo bajo por parte de la población para la adquisición de alimentos nutritivos. Este porcentaje de la población cuyo ingreso no le permite consumir alimentos saludables, aunado a aquellas personas que tienen el poder de compra, pero una mala educación alimentaria (en la región Norte de México se destina una mayor proporción de gasto a la compra de alimentos no sanos, mientras que en la región Sur se destina una mayor proporción a alimentos saludables, [4] ) da como resultado una población vulnerable en cuanto a salud se refiere. Es importante remarcar que gran parte de la población clasifica a los alimentos industrializados dentro de la categoría de alimentos no saludables pues los asocia con el incremento en las prevalencias de sobrepeso y obesidad. En México el porcentaje nacional de personas con sobrepeso (39.5%) u obesidad (35.3%) es más alto que el promedio mundial [4]. Estas dos condiciones generan en la población la presencia de enfermedades crónico-degenerativas como la hipertensión arterial (en la República Mexicana a mayor edad de los individuos, la prevalencia de hipertensión incrementa siendo más frecuente en los hombres que en las mujeres) y diabetes mellitus (la Federación Internacional de Diabetes predice que, en México, habrá nueve millones de personas con diabetes para el 2025), elevando el riego de mortalidad en los habitantes que las padecen [4].

De acuerdo con el Instituto de Información Estadística y Geográfica de Jalisco (IIEG), en el 2021 la obesidad, la hipertensión arterial y la diabetes mellitus tipo II ocuparon el 7mo, 8avo y 10mo lugar de muertes, respectivamente, en el estado de Jalisco solo porque las enfermedades respiratorias agudas, el COVID y los padecimientos agudos derivados de éste, fueron los primeros lugares (Figura 3) [5].



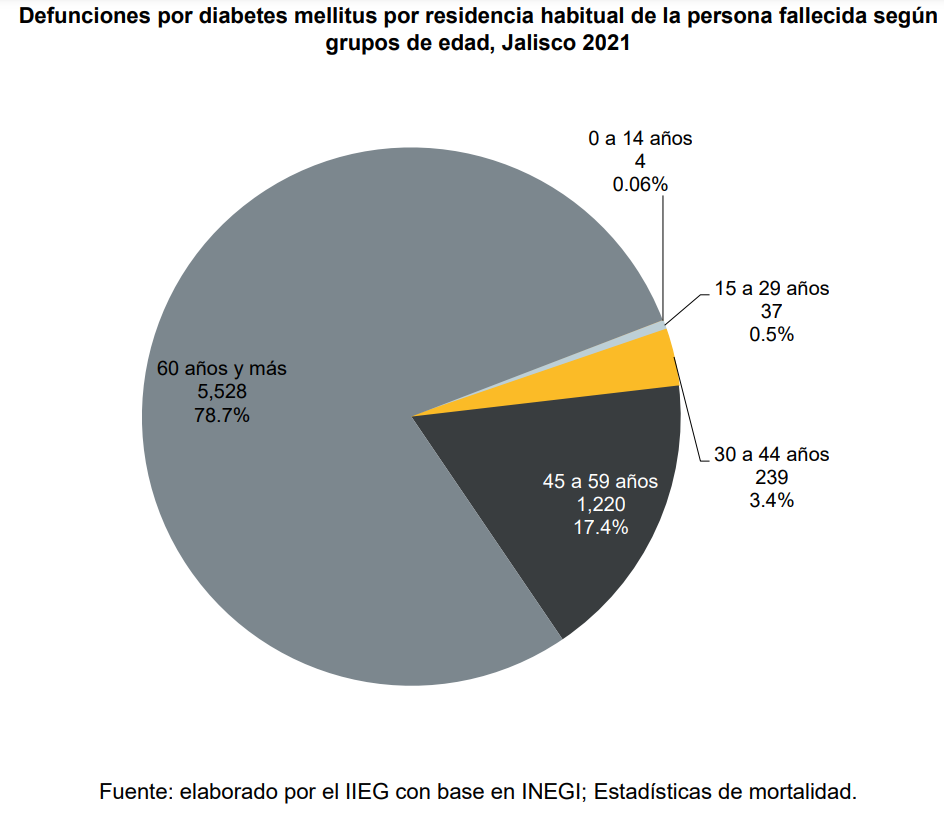
**Figura 3. Veinte principales causas de enfermedad en Jalisco durante el 2021 [6].**

Datos del INEGI (Instituto Nacional de Estadística y Geografía) reportados por el IIEG, entre 2019 y 2021, la tercera causa de mortalidad entre los habitantes del estado de Jalisco fue la diabetes mellitus (Figura 4). En el 2021 el 78.7% (5,528) de los habitantes del estado fallecidos por esta causa, fueron adultos mayores de 60 años y más; el 17.4% (1,220) estaban en un rango de edad entre 45 y 59 años y el 3.4% (239) entre 30 a 44. Cabe resaltar que, para ese año, la diabetes cobró la vida de 4 niños de entre 0 y 14 años (0.06%) y la de 37 jóvenes de entre 15 y 29 años (0.5%) (Figura 5) [5].



**Figura 4. Principales causas de mortalidad por residencia habitual y sexo de la persona fallecida en Jalisco entre el 2019 y 2021 [6].**

Las estadísticas para el 2021 mostraron que hubo mayor número de muertes de hombres que de mujeres para esta enfermedad, por grupos de edad: en el grupo de 15 a 29 años murieron por diabetes 24 hombres y 13 mujeres y en el de 30 a 44 años, fueron 138 hombres y 101 mujeres; para el grupo de entre los 45 y 59 años, fallecieron 712 hombres y 508 mujeres y entre las personas adultas de 60 años y más, murieron 2 mil 821 hombres y 2 mil 707 mujeres por causa de la diabetes mellitus [6].



**Figura 5. Defunciones por diabetes mellitus por residencia habitual en el estado de Jalisco de la persona fallecida según grupos de edad, en el 2021 [6].**

Claramente la población mexicana necesita una mejor educación alimentaria para prevenir las enfermedades crónico-degenerativas, pero también necesita mejores estrategias para disminuir la inseguridad alimentaria y el desperdicio de alimentos. Uno de los aspectos que ha contribuido a mantener estas problemáticas nacionales es el diseño y procesamiento de alimentos que no satisfacen todas las necesidades del consumidor meta, no solo en cuanto a sus características sensoriales, sino también en cuanto a ingredientes saludables, sustentables, de bajo costo y cero desperdicio; se requiere un cambio en la producción de alimentos que permita romper el círculo vicioso entre estos desafíos pues la afectación económica normalmente se traslada al consumidor quien recae en la insuficiencia de poder adquisitivo para alimentarse y cuidar de su salud, generando nuevamente inseguridad alimentaria y por ende enfermedades. Por lo tanto, es necesario un enfoque interdisciplinario para aportar soluciones innovadoras a estas grandes problemáticas que aquejan al mundo y en particular a México.

## 1.4. Planeación de alternativa(s)

Se hizo este plan de actividades del PAP junto con la Dra. Raquel Zúñiga que es para el apoyo al Ing. Daniel Hernández Mota y la alumna de maestría la Lic. Patricia Muñoz, estas actividades se harán a lo largo del verano, aunque se estipulo que algunas de no se podrían completar por el corto tiempo y la complicidad que tienen, por lo que se tomó la decisión de trabajar más con los principales modelos que fueron los que creó el Ing. Daniel Hernández Mota.

Para esto se investigó acerca de cómo se pudiera manejar los datos que se usaron, estos al ser palabras no se podían utilizar directamente en el modelo por lo que se transformaron las palabras a números por medio de diferentes técnicas lo que llevó a la generación de vectores, lo que se pidió para el mejoramiento de esos vectores fue el ver como sus datos se podrían manejar de mejor manera ya que en el modelo inicial estos se utilizaban con el promedio por lo que se investigó una mejor manera y se encontró lo siguiente, la diferencia por medio del promedio es un método que representa 𝑆 = (𝑤1, 𝑤2,…,𝑤𝑛), tomando la diferencia de todos los vectores y dividiendo el total por el número de palabras que se encuentran en estos vectores 𝑛.

Eso fue lo que se usó al inicio con el primer modelo pero en el mismo documento encontre una manera diferente de usar los vectores que fue de la siguiente manera, en este método todas las palabras dentro del vector se suman, esto se usa para escalar las suma de los vectores para que su suma al final sea igual a 1.

Al final al ver la diferencia y al leer más en el documento se puede observar que los dos son muy buenos para el manejo de las palabras en los vectores solo que el promedio es una mejor manera cuando en los vectores se están usando enunciados y palabras en conjunto y la suma es mejor cuando se tratan de palabras solamente, como los vectores que utilizamos solo contienen palabras y no enunciados se concluyó que se debería intentar el utilizar la suma para el procesamiento de los vectores [7].

También se necesitó investigar acerca de algoritmos genéticos ya que la implementación del modelo significaba que pudiéramos generar nuevos alimentos, esto se haría con el algoritmo genético implementando el modelo para ver si los nuevos productos son similares al producto que se busca. Para esto se investigó cómo funcionan los algoritmos genéticos y cómo implementarlos en Python, para esto se investigó que los algoritmos genéticos son, heurístico, bioinspirado y su función principal son las búsquedas y la optimización numérica [8].

La tabla está dividida en Nombre de la actividad son las actividades que se fueron elaborando a lo largo de la semana, Recursos lo que se necesito para poder realizar las actividades previstas, Tiempo la duración de las actividades ya sea en días o semanas, y 8 semanas, duración del verano.

| **Nombre de la actividad** | Recursos | Tiempo (días/semanas) | Semana 1 | Semana 2 | Semana 3 | Semana 4 | Semana 5 | Semana 6 | Semana 7 | Semana 8 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Familiarizarse con las bases de datos que se utilizaron. | CPU. | 1 |  |  |  |  |  |  |  |  |
| Bajar los scripts e implementarlos en el CPU personal. | GIT, CPU. | 1 |  |  |  |  |  |  |  |  |
| Entender el funcionamiento de los algoritmos para empezar a trabajar. | SUP, CPU. | 1 |  |  |  |  |  |  |  |  |
| Hacer pruebas para ver los casos triviales. | Py, CHM,SUP. | 2 |  |  |  |  |  |  |  |  |
| Investigar sobre diferentes maneras de mejorar el modelo. | TF, OV, IT. | 2 |  |  |  |  |  |  |  |  |
| Implementar lo investigado para la generación de un nuevo modelo. | GIT, Py, CHM. | 1 |  |  |  |  |  |  |  |  |
| Investigar y entender el funcionamiento del algoritmo genético usado y los algoritmos genéticos. | PXT,IT, LB, ALG. | 2 |  |  |  |  |  |  |  |  |
| Implementar el algoritmo genético de manera correcta. | Py, GIT, SUP. | 2 |  |  |  |  |  |  |  |  |
| Hacer una propuesta con los resultados dados por el algoritmo genético. | DCX. | 1 |  |  |  |  |  |  |  |  |
| Subir todo a una plataforma de internet donde se puedan ver los trabajos. | GIT |  |  |  |  |  |  |  |  |  |

| Recursos utilizados | |
| --- | --- |
| ***Abreviatura*** | Significado |
| CPU | Computadora en la que se estará trabajando. |
| GIT | GitHub página de repositorios. |
| SUP | Supervisor del PAP. |
| Py | Python, lenguaje de programación. |
| CHM | IDE o aplicación llamada PyCharm |
| TF | Página de internet llamada Tensor Flow. |
| OV | Página de internet llamada Stack Overflow |
| IT | Internet en general. |
| PXT | Presentaciones sobre los temas. |
| LB | Libros sobre los temas que se buscan. |
| ALG | Algoritmos, ya sean públicos para su estudio o proporcionados a mi |

## 1.5. Desarrollo de la propuesta de mejora

Durante el desarrollo del proyecto se pidieron varias cosas por lo que se empezará por los primeros objetivos logrados que fueron las diferentes maneras en las que se puede mejorar el modelo para esto se hizo una investigación que ya se mencionó antes y se usó para el mejoramiento del modelo, se utilizó la suma de los vectores para el mejorar el modelo lo que dio como resultado las siguientes métricas que se utilizaron para evaluar el modelo (Figura 7).

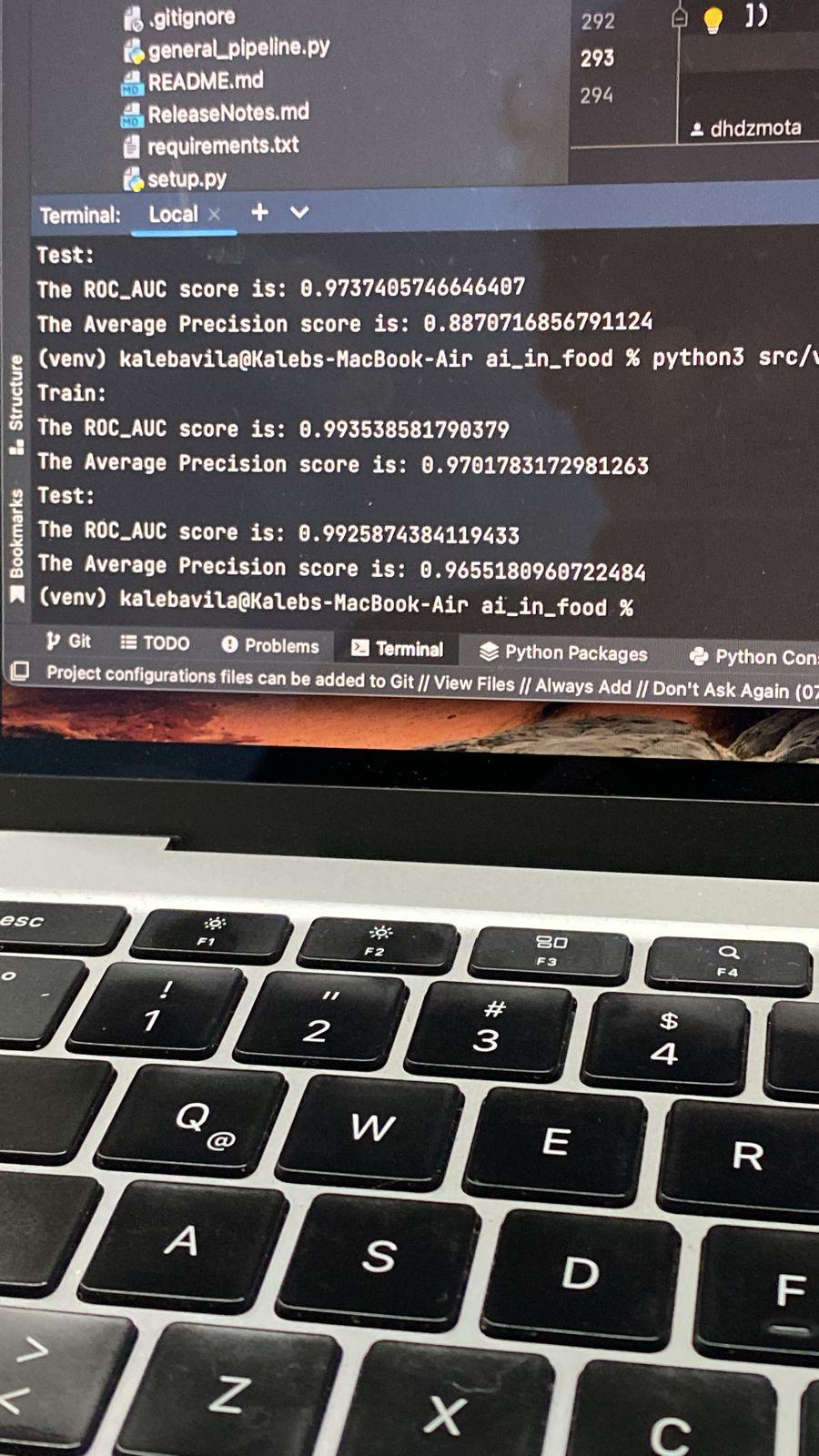


Figura 7. Se ven las evaluaciones que se obtuvieron del modelo nuevo que se generó.

Las métricas significan lo siguiente, el ROC AUC es el área bajo la curva lo que esta métrica representa es que tan bien el modelo puede discriminar entre instancias positivas y negativas, esto principalmente se utiliza para modelos de clasificación como el que se desarrolló [9].

The Average Precision es muy parecido al ROC AUC la única diferencia es que este al evaluar entre las instancias toma como peso para evaluar la instancia anterior, por lo que esta métrica es mejor para poder evaluar el modelo que se generó [8]. Este nuevo modelo que se generó fue mejor que el primero que se utilizó ya que el primer modelo usado sus métricas tenían valores al menos de 9% menos por lo que si este nuevo modelo funciona de manera correcta podría ser una gran mejora a con el que se tenía antes.

Para generar ese modelo se modificó parte del script de Python, se modificó lo siguiente (Figura 8).

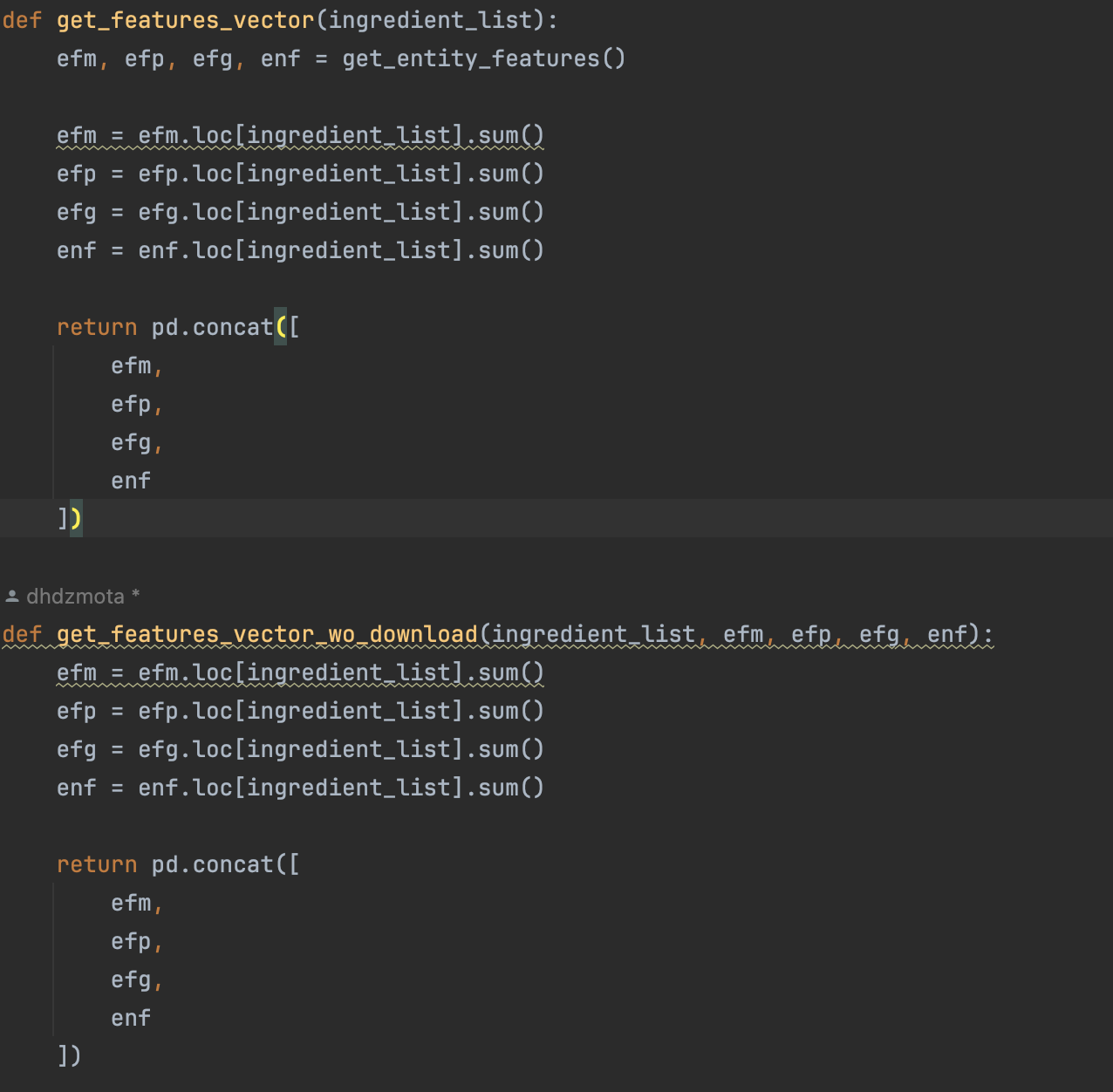


Figura 8. Se muestra el código en la parte donde se usó la suma para el manejo de los vectores.

Esos bloques de código son los que se modificaron para que en vez de que los vectores se generarán con el promedio de estos se generará mejor con la suma, dando así el nuevo modelo generado.

Después de eso se empezó a investigar acerca de los algoritmos genéticos para poder implementar ese modelo en nuevo algoritmo a parte del que ya está implementado, el nuevo algoritmo sigue en construcción pero el modelo se probó con el algoritmo ya implementado y arrojó estos resultados (Figura 9).

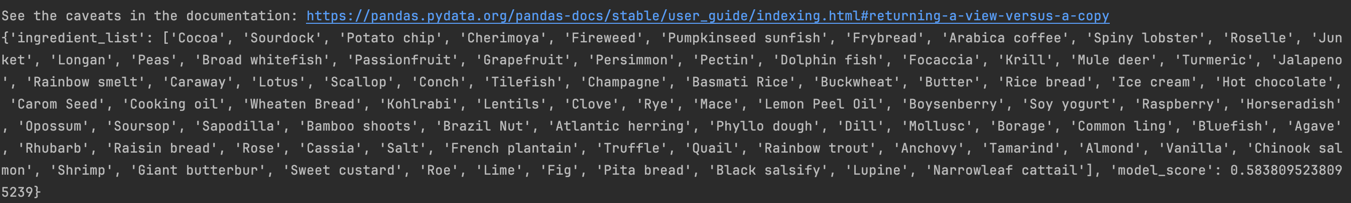


Figura 9. Lista de ingredientes para la generación de leche con una precisión del modelo de 58%

Esto generó una lista de ingredientes con los cuales se puede generar una nueva bebida parecida a la leche con un score del modelo de 58% algo que parece ser un poco bajo pero con el primer modelo el score llegó a 53% por lo que sí aumentó 5%, esto puede parecer malo pero ese score del modelo se mide de diferente manera a como las métricas este se mide de esta manera (Figura 10).

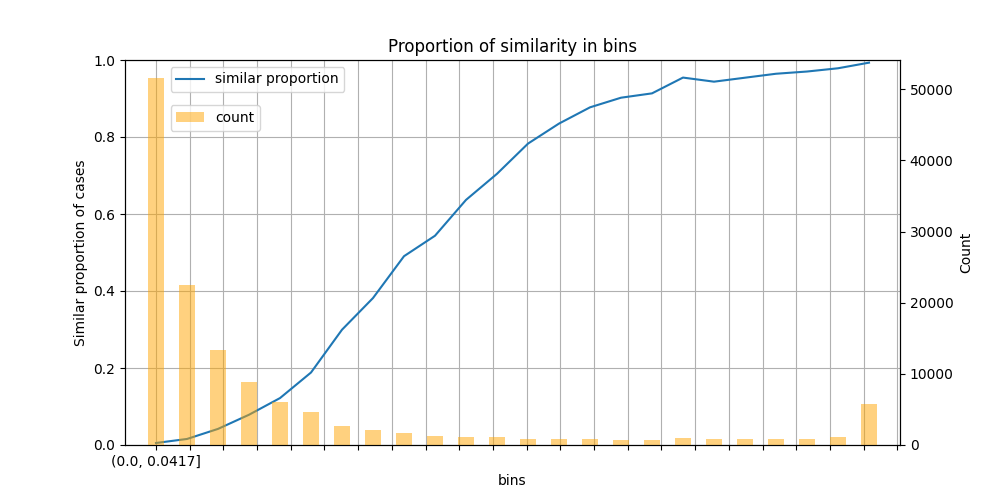


Figura 10. Es la proporción de similaridad que hay en la precisión del modelo para poder observar qué tan similar es la lista de ingredientes con el alimento a generar.

El score que se nos da el porcentaje de la izquierda es que tan similar son y lo de abajo es el score del modelo en intervalos, el score que dio con el modelo fue de 58% por lo que es un poco más de la mitad de dónde están los bins, lo que nos da una similitud con el producto a generar de más del 80% por lo que sí fue una buena generación de ingredientes los que nos dio.

Después de eso se incluyó a el algoritmo genético una nueva mutación para observar si este podría mejorar su capacidad de dar mejor calidad en las listas de ingredientes que se generan, con este nuevo algoritmo se obtuvieron varias listas de ingredientes y se compararon con unas listas generadas con el otro algoritmo genético y se obtuvieron dos nuevas propuestas para elaborar un nuevo producto parecido a la leche (Figura 11).

****

**Figura 11. Lista de ingredientes generados como propuesta para la elaboración de un producto parecido a la leche.**

Se seguirá observando y generando nuevas listas de ingredientes para ver si alguna de estas puede aumentar más la precisión del modelo y analizarla con las propuestas de ingredientes que se están usando para la elaboración de la leche.

## 1.6. Valoración de productos, resultados e impactos

Al final los productos que se desarrollaron fueron el modelo y la lista de ingredientes, estos teniendo sus respectivas métricas para su valoración, estadísticamente las métricas para valorar algo pueden tener un gran peso o ninguno ya es cuestión del investigador o la persona encargada observar y analizar si estas métricas ayudan o no, conforme a eso en lo que se entregó las métricas respectivamente que se usaron dieron para el nuevo modelo y las nuevas listas un porcentaje de 96% en el modelo respectivamente, 53.7% y 54.8% en las listas generadas en la precisión del modelo, en el primer modelo generado por Ing. Daniel Hernández Mota la métrica del modelo fue de 86% por lo que se puede decir que el uso de la suma para los vectores fue buena idea, aunque todavía se tiene que observar más en ese punto, y en la lista el score que dio fue de 53% algo muy parecidos a los otros que se generaron lo que puede decir que el agregar más mutaciones al algoritmo genético no ayudó bastante.

Todavía se puede mejorar las cuestiones del modelo, de si en vez de usar la suma de vectores ver si hay otro tipo de métodos para su manejo, el quitar o agregar mutaciones al código genético como entre muchas otras, lo que al final sí se logró fue el mejorar un poco el modelo por medio de investigación y pruebas para poder llegar al objetivo, esto también pudo ser de ayuda a otras actividades como en el otro modelo que se estaba generando también se estaba usando otro método diferente a la suma de vectores y con lo que se investigó se planteó el si se hacía un cambio o no a ese modelo.

Los resultados dados fueron de mucha ayuda a la solución de la problemática, esto ya que al estar investigando y poniendo a prueba todo lo que ya se había hecho dentro del proyecto se encontraron algunos errores dentro del código así como en las cosas básicas dentro de este, como el ayudar a explicar cómo entender el código para que nuevos usuarios o nuevas personas puedan hacer uso de este, el poner a prueba todo para ver su funcionamiento correcto así como el reportar los errores encontrados, los resultados también ayudaron a que se pueda obtener un mejor producto final al dar un listado un poco mejor de los ingredientes así como varios ingredientes endémicos.

Algo que faltó por hacer es el hacer más pruebas con diferentes métodos, así como el hacer un modelo nuevo desde cero, pero por falta de tiempo se prefirió ayudar en los modelos que ya se estaban desarrollando.

## 1.7. Bibliografía y otros recursos

[1] “Recetas”. México. Accedido el 28 de junio de 2024. [En línea]. Disponible: <https://notco.com/mx/recetas>

[2] “word2vec | Text | TensorFlow”. Tensor Flow. Accedido el 28 de junio de 2024. [En línea]. Disponible: <https://www.tensorflow.org/text/tutorials/word2vec>

[3] “Se desperdician Más de diez mil toneladas de alimentos cada año en México”. gob.mx. Accedido el 28 de junio de 2024. [En línea]. Disponible: <https://www.gob.mx/bienestar/prensa/se-desperdician-mas-de-diez-mil-toneladas-de-alimentos-cada-ano-en-mexico>

[4] ENSANUT. (2018). Encuesta Nacional de Salud y Nutrición 2018-2019 Resultados nacionales.

[5] IIEG. (2022). Día Mundial de la Diabetes 2022.

[6] IIEG. (2021). Enfermedades de la población de Jalisco en 2021.

[7] Dilawar, N.; Majeed, H.; Beg, M.O.; Ejaz, N.; Muhammad, K.; Mehmood, I.; Nam, Y. Understanding Citizen Issues through Reviews: A Step towards Data Informed Planning in Smart Cities. *Appl. Sci.* **2018**, *8*, 1589. <https://doi.org/10.3390/app8091589>

[8] D. C. Lozano. “Clase Métodos de inteligencia Artificial”. Algoritmos genéticos.

[9] “average\_precision\_score”. scikit-learn. Accedido el 28 de junio de 2024. [En línea]. Disponible: <https://scikit-learn.org/stable/modules/generated/sklearn.metrics.average_precision_score.html>

[10] “How to explain the ROC AUC score and ROC curve?” Evidently AI - AI Observability and ML Monitoring Platform. Accedido el 28 de junio de 2024. [En línea]. Disponible: <https://www.evidentlyai.com/classification-metrics/explain-roc-curve#:~:text=The%20ROC%20AUC%20score%20is%20the%20area%20under%20the%20ROC,and%201%20indicates%20perfect%20performance.>

## 1.8. Anexos generales

## Productos

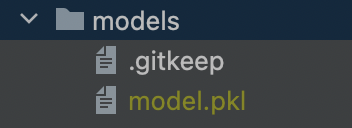
Se entregarán tres productos que son lo que ayuda más a la solución del problema, el modelo en un archivo pickle para que se pueda usar por cualquier persona y puedan hacer sus propios análisis así como el generar nuevas listas para nuevos productos, y el repositorio donde se encuentran los scripts que se usaron a lo largo del verano para que cualquier persona de la institución pueda tener acceso a este y pueda hacer uso de todo lo que se trabajó.

El modelo se elaboró usando los scripts dentro del repositorio, esto tomando los datos que se necesitan, sus procesamientos correspondientes y al final el entrenamiento de este.

El repositorio fue elaborado a lo largo del proyecto y es donde se fueron subiendo las actualizaciones del proyecto, así como las correcciones de los errores que se fueron dando.

Modelo.

| Nombre y código del PAP | 4G03 Programa de apoyo a la investigación I |
| --- | --- |
| Nombre del proyecto | Investigación de formas de procesar palabras y optimización del código generado para el mejoramiento de modelos alimenticios generados por (IA). |
| Descripción (qué es, para quién se realizó y para qué es): | Es un modelo guardado en un archivo pickle que sirve para que no tengas que volver a entrar un modelo desde cero y lo puedas usar en diferentes ambientes. |
| Autores: | Kaleb Azael Avila Padilla y Daniel Hérnandez Mota |



Otro.

| Nombre y código del PAP | 4G03 Programa de apoyo a la investigación I |
| --- | --- |
| Nombre del proyecto | Investigación de formas de procesar palabras y optimización del código generado para el mejoramiento de modelos alimenticios generados por (IA). |
| Descripción (qué es, para quién se realizó y para qué es): | Un repositorio en internet donde se guardan los scripts del proyecto. |
| Autores: | Kaleb Azael Avila Padilla y Daniel Hérnandez Mota |

<https://github.com/KalebAvila/ingenieria_en_alimentos/tree/develop>

## 3. Reflexión crítica y ética de la experiencia

El RPAP tiene también como propósito documentar la reflexión sobre los aprendizajes en sus múltiples dimensiones, las implicaciones éticas y los aportes sociales del proyecto para compartir una comprensión crítica y amplia de las problemáticas en las que se intervino.

## 3.1 Sensibilización ante las realidades

Al inicio del proyecto la verdad no pensaba que el alimento en sí era un gran problema y más el como se planteó como problemática ya que creía que los alimentos que eran hechos a base de otro como el pollo o carne de plantas era más para gente privilegiada y nunca se me ocurrió que gente lo comiera por saludable o sus nutrientes, después de investigar todo eso me hizo dar cuenta que puede ser algo complicado para todos y que el hacer nuevos alimentos con ingredientes endémicos de México podría ser una gran solución por que eso haría que fuera más barato y accesible el conseguir esos alimentos.

Conforme avanzaba con mis actividades empece a pensar el como puede que las nuevas tecnologías ayudan mucho a la gente y puedan ser de gran ayuda a problemas que se encuentran desde hace mucho, pero también me deja a pensar que puede que personas usen estas nuevas tecnologías o investigaciones para hacer cosas que no son éticas o aprovecharse de eso para sacar un beneficio propio sin importarle dañar a los demás, por lo que si en un futuro trabajo en algo parecido tener en cuenta el propósito de lo que se está haciendo.

## 3.2 Aprendizajes logrados

Lo que más se me complicó durante todo el proyecto fue el entender cómo funcionaba este ya que yo al no estar presente cuando se desarrolló tenía que aprender todo desde cero y ver cómo funcionan los scripts y los códigos que ya se habían desarrollado anteriormente, justo con todo eso lo complicado fue el modificarlo y que funcionara ya que al ser algo que no fue creado por mi cuenta al moverlo hacia que fallaba algo desde muy atrás o algo que estaba implementado desde el inicio y pues el arreglarlo y hacer que funcionara correctamente fue algo muy difícil.

También algo que se me complico mucho fue el investigar nuevas técnicas y nuevas prácticas que nunca había trabajado, como lo fue el algoritmo genético y el manejo de tokens en las palabras, al nunca haber trabajado con eso tuve que leer mucho e investigar sobre esos temas como a su vez ponerlos a prueba para saber cómo funcionaban.

Al final lo que más se me complico fueron los tiempos ya que al tener otra materia y al tener poco tiempo para entregas el esperar horas para que el código pudiera funcionar pues era algo frustrante ya que en muchos scripts estos tardaban una hora o más por tener demasiados datos, lo bueno fue que con todo esto aprendí mucho a tomar mis tiempos y ser paciente en los trabajos, así como el leer informes y papers de mi carrera, pero lo que más aprendí fueron las nuevas formas de procesamientos de datos y el poder implementar el script de manera eficiente con nuevas cosas.

*[*En este apartado escriben de manera personal las competencias disciplinares, sociales y universitarias que lograste desarrollar durante el trabajo en el PAP, especifica en qué momentos sentiste el reto y cómo fue que respondiste, dándote cuenta de que eres capaz de hacerlo bien.

También especifica los aprendizajes que lograste al ser parte de este proyecto, en el ámbito profesional, social y personal.*]*

3.3 Inventario de competencias Inicial (ingreso del PAP) e Inventario de competencias Final (salida al PAP).

**Tabla de inventario de competencias al inicio del PAP**

**Nombre del estudiante:** Kaleb Azael Avila Padilla

| Categoriza las competencias en conocimientos, habilidades y actitudes.    Escribe la o las evidencias de cada competencia y su relevancia. | **Competencia** | | **Evidencia** | **Relevancia/Fortaleza\*** | **Competencias**  **Nuevas** | **Competencias**  **Potenciadas.** |
| --- | --- | --- | --- | --- | --- | --- |
| **Conocimientos** | Conocimiento de varios modelos matemáticos, así como su implementación en código. | Varios proyectos a lo largo de la carrera, muchos con redes neuronales y otros con modelos de ML, como SVM, K-means, entre otros. | Puedo aportar al proyecto gran ayuda ya que se está usando un modelo matemático para la investigación. | Aprendí de algoritmos genéticos y su implementación en el lenguaje que se solicitó.  Aprendí a tener más capacidad de amoldarse a nuevos proyectos o el adaptarme rápido a estos a pesar de no estar presente al inicio de este.  El saber buscar y leer investigaciones sobre nuevos temas para poder ponerlos en práctica o como referencia. | Me ayudó a mejorar mis conocimientos que ya tenía sobre modelos matemáticos y programación.  Me ayudó a poder adaptarme mejor a las cosas y ver como en un futuro poder tomar estas enseñanzas para proyectos futuros.  Me ayudó mucho a tener más capacitado mi análisis matemático y el saber como obtener la información relevante de las investigaciones. |
| Varios lenguajes de programación y la base de estos | Las clases impartidas a lo largo de mi carrera, así como las clases optativas que lleve de estas, como un certificado en JAVA Script. | El ayudar a hacer más accesible o fácil de usar el código, así como también el optimizarlo o ayudar en errores que pueda haber en este. |
|  |  |  |
|  |  |  |
|  |  |  |
| **Habilidades** | Análisis matemático como buena capacidad de trabajo. | El analizar varios research a lo largo de la carrera, como también las clases de matemáticas impartidas como estadística bayesiana. | El poder ayudar a complementar el paper así como el dar nuevas opciones u opiniones de cómo se podría mejorar el análisis matemático o el uso de ciertos conceptos matemáticos para el modelo. |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
| **Actitudes** | Poder adaptarme a la forma de trabajo y el poder aprender rápido los temas nuevos o necesarios para trabajar. | Puedo trabajar de la manera que se me solicite ya que puedo adaptarme rápido a los nuevos o viejos métodos que sean necesarios, así como también poder estudiar y aprender lo que se necesite para el proyecto. | Apoyar a los del equipo cuando sea y de una manera proactiva para que el proyecto no se estanque y pueda ser terminado a tiempo y de manera correcta. |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |