**Fase 4: Transición de la formulación de ideas a los diseños preliminares**

**Ideas descartadas:**

Pata generar Frequent Itemsets:

Como en el problema se extraen relativamente pocos itemsets comparados con todos los ítems dados, entonces su funcionalidad deja de ser eficaz ya que se espera que haya una gran densidad de transacciones con respecto a la barrera de los frecuentes.

Para el análisis de las transacciones de Allers:

El método de Support Vector Machine utiliza la técnica del aprendizaje supervisado, lo cual se limita para desarrolladores estudiantes debido al poco tiempo que se tiene, requeriría hacer etiquetas a conjuntos de datos y hacer un proceso mucho más elaborado, además, existen otros métodos ligados a la predicción y con menores recursos.

Presentación de la información en la GUI:

Aunque uno de los objetivos es separar grupos de clientes a los cuales se les podría hacer descuentos u otro tipo de estrategias de marketing, no se considera relevante observar la información de cada cliente de manera individual o ver sus posibles compras mediante cuadros o gráficos. Sin embargo, podría hacerse por grupos de clientes en el transcurso de la realización del proyecto.

Análisis de los clientes:

Por razone similares a la idea descartada en análisis de transacciones, el SVM resulta ser una alternativa poco probable tanto para transacciones como para clientes, en general, la técnica del aprendizaje supervisado requiere de más detalle para su elaboración.

**Diseños preliminares:**

Para generar Frequent ItemSets:

1. El algoritmo realiza una búsqueda de profundidad para saber cuáles no deberían hacer parte de los frecuentes y así descartarlos inmediatamente.
2. Usar un método eficiente que tenga una representación robusta de las transacciones por medio de un árbol, Growth Algoritm es útil para extraer de grandes volúmenes de transacciones para su posterior análisis

Análisis de las transacciones de Allers:

1. La aplicación contiene métodos relacionados con análisis de asociación. Esto es, la aplicación ejecuta procedimientos que aplican el análisis de asociación teniendo en cuenta los métodos predefinidos, las diferentes técnicas y variabilidad en las maneras de recolectar y analizar los datos.
2. Un prototipo del aplicativo que contenga métodos como medidas de tendencia central, anovas, estudio de distribuciones, aplicar probabilidades y concluir con lo que se obtiene usando estadística descriptiva.
3. La aplicación se implementa con matrices multidimensionales para asociar a cada ítem con la transacción y a su vez con los clientes, discretizando algunos atributos para hacer respectivas aproximaciones, en esta aplicación se dejan de utilizar filas y/o columnas para agrupar los objetos del problema.

Presentación de la información en la GUI:

1. El programa muestra la información por medio de gráficos de barras, dispersión y tortas, apoyadas por datos numéricos representando dispersión y medidas de tendencia central. Cada gráfica representa: los porcentajes de aparición de cada ítem, las variaciones en las ventas de cada producto debido a las transacciones y la comparación de conjuntos de ítems.
2. Se separa la información de transacciones por mes, esto para facilitar las predicciones y notar más fácilmente el crecimiento o decrecimiento en las ventas de ciertos productos. Esta vez se muestran las reglas extraídas por mes y soportes.

Análisis de los clientes:

1. El programa se implementa con una gran variedad de métricas sobre las transacciones, con las cuales se va a poder medir la similitud entre clientes, y así poder encontrar productos que unos clientes no compran pero que podrían comprar dado su similitud con otros clientes que si los compran.
2. La aplicación contiene un método de agrupamiento denominado k-means, que ayuda en gran medida a la división de los clientes basados en sus transacciones y al trato de los mismos para realizar análisis más específicos.