

Conclusiones del Análisis Experimental de Consumo en Telecomunicaciones

En este trabajo se realizó un análisis experimental del **comportamiento de consumo de usuarios de servicios de telecomunicaciones**, considerando voz, SMS y datos móviles, con el objetivo de identificar **patrones de uso**, clasificar **usuarios según su relación de consumo** y evaluar **cómo dichos patrones varían en función del tiempo**.

1. Proceso metodológico seguido

1.1. 1. Preparación y limpieza de datos

El análisis comenzó con la **conversión de las variables temporales** (`START_DATE`, `END_DATE`) a formato datetime y el cálculo de la **duración de cada sesión** en horas. Posteriormente, se eliminaron **outliers previamente detectados** mediante DBSCAN, con el fin de evitar que comportamientos extremos distorsionaran el proceso de clustering.

Este paso permitió trabajar con un conjunto de datos más representativo del comportamiento típico de los usuarios.

1.2. 2. Construcción del perfil de usuario

Para cada usuario (`OBJ_ID`), se agregaron los consumos por servicio y se transformaron en **proporciones relativas**:

- Proporción de consumo en voz
- Proporción de consumo en SMS
- Proporción de consumo en datos

Este enfoque permitió:

- Comparar usuarios independientemente del volumen absoluto de consumo

- Capturar la **relación entre servicios**, que era el objetivo central del análisis

La suma de las proporciones por usuario es igual a 1, garantizando una representación normalizada del comportamiento.

1.3. 3. Elección de categorías de usuario

Inicialmente se planteó definir **categorías manuales** (por ejemplo, “usuario mixto”, “usuario balanceado”), estableciendo márgenes porcentuales entre servicios. Sin embargo, los resultados obtenidos mediante clustering mostraron que:

- Los usuarios se agrupan de forma **natural y extremadamente separada**
- Los centroides están **claramente dominados por un único servicio**
- No existen zonas intermedias significativas entre clusters

Por esta razón, **no tuvo sentido forzar categorías artificiales**, ya que los propios datos revelaron perfiles de consumo bien definidos. La categorización final se realizó de forma **data-driven**, usando K-Means sobre las proporciones de consumo.

1.4. 4. Clustering de usuarios

Se utilizó **K-Means** con tres clusters, justificados por:

- El número de servicios analizados (voz, SMS, datos)
- La clara separación observada entre los grupos
- La interpretabilidad directa de los centroides

Los clusters resultantes representan:

- Usuarios dominados por consumo de datos
- Usuarios dominados por consumo de SMS
- Usuarios dominados por consumo de voz

La alta pureza de los centroides confirmó la robustez del modelo y la claridad de los patrones presentes en los datos.

1.5. 5. Análisis temporal por franjas horarias

El dataset se segmentó en distintas franjas horarias:

- Hora Sueño (00–05)
- Despertar (05–08)
- Hora Laboral (08–12)

Para cada franja:

- Se reconstruyeron los perfiles de usuario
- Se aplicó nuevamente el clustering
- Se analizaron centroides y distribuciones relativas

Este enfoque permitió evaluar la **estabilidad temporal de los perfiles de consumo**.

2. Resultados principales

2.1. Especialización del consumo

En todas las franjas horarias, los usuarios muestran un **alto grado de especialización**, concentrando casi la totalidad de su consumo en un único servicio. Esto se refleja en centroides con proporciones cercanas al 100 % para uno de los servicios y prácticamente nulas para los demás.

2.2. Ausencia de usuarios mixtos

Contrario a lo esperado inicialmente, **no se identificaron perfiles de consumo balanceado**. Los usuarios no combinan de forma significativa voz, SMS y datos dentro de las ventanas temporales analizadas.

Este resultado indica que:

- El comportamiento de consumo es altamente segmentado
 - Los servicios no se utilizan de forma simultánea o complementaria en el corto plazo
-

2.3. Influencia de la franja horaria

La franja horaria **no altera la estructura de los perfiles**, pero sí afecta su **prevalencia**:

- El consumo de datos domina en todas las franjas
- La proporción de usuarios de voz y SMS varía según el momento del día

Esto sugiere que el tiempo influye en **cuántos usuarios** pertenecen a cada tipo, pero no en **qué tipos existen**.

3. Conclusión general

El análisis demuestra que el consumo de servicios de telecomunicaciones presenta **patrones de uso altamente especializados**, estables en el tiempo y claramente separables mediante técnicas de clustering no supervisado.

La experimentación evidenció que:

- La clasificación basada en proporciones es más adecuada que el uso de valores absolutos
- No es necesario definir categorías manuales cuando los datos muestran una separación natural
- El comportamiento de los usuarios es consistente y predecible dentro de las franjas analizadas

Este estudio sienta una base sólida para análisis futuros, como:

- Evolución del usuario entre perfiles a largo plazo
- Detección de cambios de comportamiento ante eventos externos
- Segmentación avanzada para optimización de servicios o gestión de fallos