



UNIVERSIDAD  
DE LA FRONTERA

# Análisis de caso Manzanar - Censo 2017 - Revisión 3

Arturo Avendaño - Alonso Rojas - Kianush Atighi-Moghaddam - Diego Vera



# Objetivos principales

- Mejorar la calidad de los datos del CENSO.
- Predecir Atributos de las manzanas.
- Obtener una precisión aceptable en el modelado.



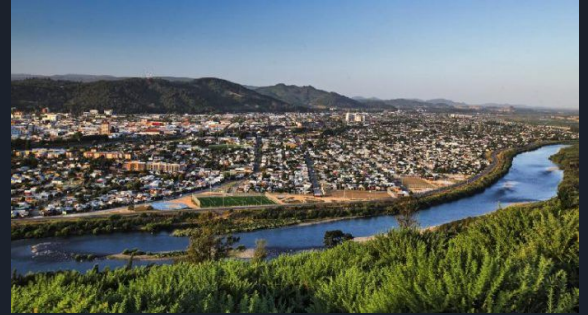
# Motivación

- Dificultad para tomar datos.
- Manzanas con datos indeterminados.
- Agilización de la toma de datos.



# Problemas abordados

1. Problema de clasificación del atributo NOM\_CAT\_ENT (Campo, Pueblo, Ciudad, Caserío, Etc.).
2. Problema de clasificación del atributo NOM\_AREA (Urbano, Rural).



Ciudad



Pueblo



# Datos

- Información más relevante obtenida de la exploración de datos:
  - Datos faltantes,
  - Estadísticas relevantes
  - Eliminación de atributos, etc

**NaN**

**Indeterminado**



# Experimentos y resultados relevantes

## Problema 1

- Elección de datos.
- Modelos de clasificación.
- Metodología.

- ORIGEN\_AGUA\_RED\_PUBLICA
- ORIGEN\_AGUA\_POZO
- ORIGEN\_AGUA\_CAMION\_ALJIBE
- ORIGEN\_AGUA\_RIO\_VERTIENTE\_ESTERO\_CANAL\_LAGO
- TIPO\_VIV\_CASA
- TIPO\_VIV\_DPTO
- TIPO\_VIV\_TRADICIONAL
- TIPO\_VIV\_CONVENTILLO
- TIPO\_VIV\_MEDIAGUA
- TIPO\_VIV\_MOVIL
- TIPO\_VIV\_OTRO
- Indice\_mat

# Experimentos y resultados relevantes

## Problema 1

### Árbol de decisión

precision	recall	f1-score	support	
Aldea	0.62	0.29	0.40	218
Asentamiento Minero	0.32	0.15	0.20	75
Asentamiento Pesquero	0.35	0.57	0.43	75
Campamento	0.00	0.00	0.00	16
Caserío	0.46	0.31	0.37	1101
Comunidad Indígena	0.61	0.02	0.04	647
Fundo-Estancia-Hacienda	0.45	0.45	0.45	1029
Otros	0.00	0.00	0.00	91
Parcela de Agrado	0.45	0.18	0.26	490
Parcela-Hijuela	0.55	0.80	0.65	3307
Veranada-Majada-Aguada	0.39	0.34	0.37	148
accuracy	0.52	0.7197		
macro avg	0.38	0.28	0.29	0.7197
weighted avg	0.51	0.52	0.47	0.7197

### KNN

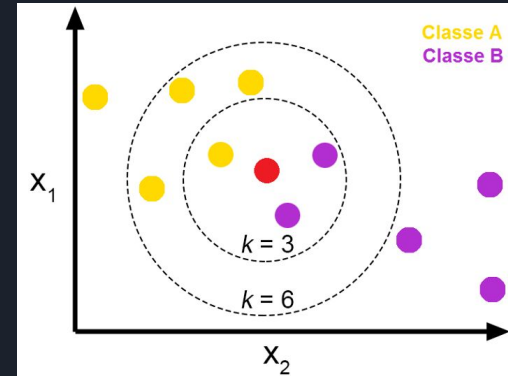
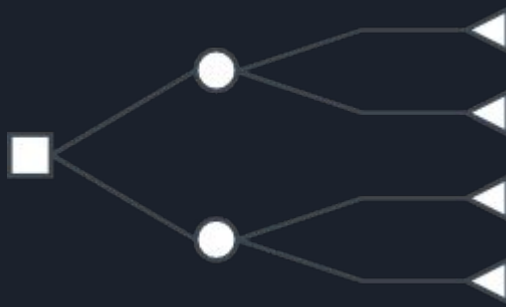
precision	recall	f1-score	support	
Aldea	0.36	0.24	0.29	218
Asentamiento Minero	0.00	0.00	0.00	75
Asentamiento Pesquero	0.32	0.11	0.16	75
Campamento	0.00	0.00	0.00	16
Caserío	0.43	0.32	0.36	1101
Comunidad Indígena	0.41	0.09	0.15	647
Fundo-Estancia-Hacienda	0.43	0.51	0.47	1029
Otros	0.25	0.27	0.26	91
Parcela de Agrado	0.56	0.21	0.31	490
Parcela-Hijuela	0.56	0.75	0.64	3307
Veranada-Majada-Aguada	0.39	0.34	0.37	148
accuracy	0.51	0.7197		
macro avg	0.34	0.26	0.27	0.7197
weighted avg	0.48	0.51	0.47	0.7197



# Experimentos y resultados relevantes

## Problema 2

- Elección de datos.
- Modelos de clasificación.
- Metodología.







# Experimentos y resultados relevantes

## Problema 2

### Árbol de decisión

precision	recall	f1-score	support	
RURAL	0.77	0.83	0.80	8576
URBANO	0.97	0.95	0.96	45574
accuracy	0.93	54150		
macro avg	0.87	0.89	0.88	54150
weighted avg	0.94	0.93	0.93	54150

### KNN

precision	recall	f1-score	support	
RURAL	0.78	0.82	0.80	8576
URBANO	0.97	0.96	0.96	45574
accuracy	0.93	54150		
macro avg	0.87	0.89	0.88	54150
weighted avg	0.94	0.93	0.94	54150

# Experimentos y resultados relevantes

## Conclusiones

- Consecuencias por no aplicar técnicas de manejo de datos necesarias (Oversampling/Undersampling).
- Calidad de los resultados.
- Observaciones.



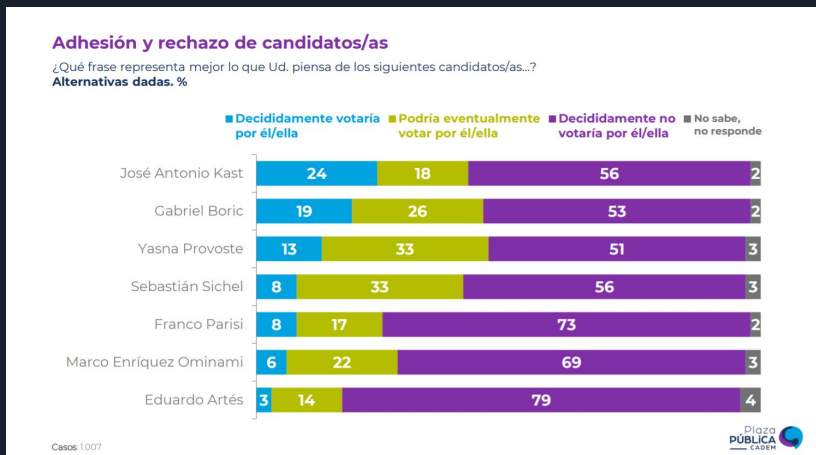
## Discusión a futuro

- En cuanto a aplicaciones a futuro que puedan surgir con los resultados y observaciones obtenidas, se encuentra el planteamiento de utilizar el atributo calculado, Índice de Materialidad como punto entre los manzanares con más urgencia de mejora y renovación de construcción.



# Discusión a futuro

- También se identifica la posibilidad de explorar este tipo de modelos para otro tipo de problemas asociados al gobierno, por ejemplo, ayudar a la encuesta CADEM a realizar predicciones de votos a lo largo del país.





# Conclusión

- Se obtuvo una precisión deplorable en el modelado de categoría.
- El estudio fue significativo para modelar el problema y plantearlo para futuros estudios con otras variables que permitan refinar el resultado.
- Se obtuvo un significativo resultado en el problema de clasificación de categoría urbana o rural.



## Link presentación

- [https://youtu.be/JpJ6P6faU\\_8](https://youtu.be/JpJ6P6faU_8)