

Descripción del problema

Innovación y Escalabilidad

- México está experimentando su **segunda sequía** registrada más fuerte en su historia.
 - Esta afecta a más del **85%** del territorio nacional.
- Las sequías se definen normalmente como una **disminución de precipitación** en una región respecto a su **índice anual**.
 - Sin embargo, estas también involucran cambios en humedad del suelo, reservas de agua, respuesta vegetal, y niveles de escorrentía
 - Más de la mitad de las **210 presas más grandes** en México están a menos del **50% de su capacidad**.
- Las sequías afectan mayormente al sector **agrícola** ya que este consume el **75%** del agua en México.
 - En la sequía del **2011**, hubo más de **8 mil millones de pesos** en pérdidas sólo en el sector **agropecuario**.
 - Esta baja producción agrícola termina afectando mayormente la **economía, nutrición, y estabilidad social** de varias **comunidades marginalizadas** en la república.
- Otras **consecuencias** significativas de las sequías incluyen
 - incremento de incendios forestales (como los más de 80 incendios forestales activos),
 - costos en el sector salud,
 - deterioro de los servicios de ecosistemas,
 - Reducción de la producción industrial.
- Las áreas **áridas y semiáridas** son más propensas a sequías, esto representa la mayor parte del **Norte y Centro de México**.
- Las sequías son generadas por **patrones oceánicos y atmosféricos** fuera de nuestro control
 - La **crisis climática**, ha incrementado la **sucesión e intensidad** de las sequías alrededor de todo mundo.
- No se es posible detener una sequía, pero sí es muy posible estar mejor prevenidos para enfrentarlas.

Soluciones Actuales

- Históricamente, el **gobierno** ha actuado de **manera reactiva** a las sequías.
 - Abasteciendo de agua a comunidades afectadas,
 - activando esquemas de aseguramiento y financiamiento para reducir los impactos económicos,
 - y limitando el consumo de agua.
 - En la sequía del **2011**, estas respuestas costaron **37 mil millones de pesos**.

- No fue hasta el **2013**, aprendiendo de la sequía pasada, que se desarrolló una herramienta federal preventiva contra las sequías. El **Programa Nacional Contra la Sequía**, conocido como PRONACOSE.
 - Este establece los **Programas de Medidas Preventivas y de Mitigación a la Sequía**,
 - Establecen **metodologías** para que **organismos locales** desarrollen y ejecuten **planes de mitigación** a las sequías a nivel de **cuenca hidrológica**.
- El **éxito** de estos programas preventivos, especialmente a nivel local, depende grandemente de tener **fácil acceso a predicciones de sequías** con tiempo de anticipación.
- **En otros países** se han desarrollado distintos **modelos muy accesibles** para predecir el suceso e intensidad de estas crisis.
 - Por ejemplo, la **Oficina Nacional de Administración Oceánica y Atmosférica de Estados Unidos** cuenta con modelos de fácil uso por página web que comunican la ocurrencia, continuación, o disipación de una sequía.
- El **gobierno de México** actualmente **solo** provee el **Monitor de Sequía en México**, el cual solo da **monitoreo de las condiciones actuales** de la sequía, no provee pronósticos especializados en sequías.

Solución

Objetivo del proyecto

- Por lo tanto, el objetivo de nuestro proyecto fue **desarrollar un modelo estadístico** en una **plataforma de fácil uso** para realizar predicciones con un mes de anticipación de las sequías en la república mexicana.
- Este modelo usa **aprendizaje automático** para realizar sus predicciones y tiene una métrica de **0.80 de área bajo la curva de precisión-exhaustividad**.

Target

- El **fácil acceso** a la predicción confiable de sequías, permitirá a los tomadores de decisiones locales **minimizar los costos** que estas presentarán en las décadas por venir.
- Grupos como políticos locales, administradores hídricos, ejidatarios, negocios privados, grupos indígenas, y principalmente agricultores podrán mejorar sus planes de adaptación a sequías.
 - Por ejemplo, un administrador de recursos hídricos se verá favorecido por el modelo de predicción local ya que podrá regular el consumo de agua en su comunidad con mayor anticipación, evitando la falta de disponibilidad de agua cuando llegue la sequía.

- Similarmente, los agricultores podrán más eficientemente administrar sus cultivos teniendo un mejor conocimiento de las condiciones hidrológicas por venir.

Fuentes de información

El corazón de este proyecto se basa primordialmente en el *Monitor de Sequía de México*. Esos datos son proporcionados por el Servicio Meteorológico Nacional (SMN) utilizando distintas redes de observación. Este conjunto de datos consiste en información municipal de cada entidad de la república mexicana y contempla datos desde el 31 de enero del 2003 hasta la actualidad¹.

Adicional a estos datos, se cuenta con una API ²llamada Dark Sky que proporciona datos meteorológicos al proporcionarle una geolocalización y una fecha en particular.

También se incorporan datos del Portal de Geoinformación 2021, del SNIB. Debido a que no tienen información de manera temporal, se obtuvieron aquellos datos que se mantienen relativamente estáticos a través de la escala temporal que se está utilizando.

En específico, se obtuvo información acerca de: cuencas, subcuencas, curvatura de perfil, curvatura del plano, curvatura del terreno, edafología, hidrología, hidrografía, hipsometría, índice topográfico de humedad y pendiente.

Explicación del modelo

Para realizar el modelo, se contempló utilizar un algoritmo de ensamble de árboles conocido como *potenciación del gradiente*. Esto hace que, a lo largo de varias iteraciones, el error de aprendizaje se vaya reduciendo y se converja a un valor de predicción. El algoritmo de potenciación del gradiente funciona de tal manera que se realiza un aprendizaje de los datos a través de la generación continua de *árboles de decisión*. Cada árbol de decisión se añade individualmente a lo largo de muchas iteraciones, ajustándose para corregir los errores de predicción de todos los árboles previamente añadidos. reduciendo el error del valor de predicción.

Distribución de la variable de respuesta

Esta gráfica nos muestra la separabilidad de la variable de respuesta. El modelo asigna un coeficiente que va del 0 al 1. En este caso el 0 significa no sequía y el 1 sequía. El color naranja está demostrando los casos de sequía y el azul la no sequía

Precision Recall Curve

La gráfica representa que el modelo llega al valor deseado, que en este caso es 0.80

Agrupación de score promedio respecto a la variable de respuesta en Jalisco

¹

²

.....

Montaje del modelo

Se busca que la experiencia del usuario en el sitio consista en resolver problemas, en este caso saber cuándo sucederá una sequía abarcando todo el territorio de la república mexicana para ayudar a los usuarios a alcanzar sus objetivos y mitigar el riesgo que conlleva este desastre natural.

- En la página principal se presenta la "Propuesta de Valor". Es decir una descripción breve de la utilidad del sitio. Es lo primero que ven los usuarios cuando ingresan para decirles exactamente lo que podemos hacer por ellos.
- Consecuentemente el usuario puede explorar en las distintas pestañas del sitio información relevante cómo para quién está dirigido el producto, por qué las sequías es una problemática de gran impacto, los beneficios al usar el modelo, los objetivos y la explicación de cómo usarlo.

En el video podemos observar a un usuario navegando en el mapa. La persona debe seleccionar el área de interés para verificar el estado de sequía y de esta manera el mapa marcará el área de sequía de un color diferente dependiendo la intensidad de la misma.

Siguientes Pasos

El desempeño del modelo es lo mínimo esperado sin embargo se puede observar el potencial predictivo que tiene a través del tiempo. Para lograr hacer esto, todavía queda trabajo por hacer.

Se requiere:

- Incorporar variables de índole global para ayudar a tener un contexto generalizado. Por ejemplo, se planea incluir variables como el fenómeno de “El niño y La niña”.
- Incorporar más fuentes de datos para complementar la información (datos meteorológicos de la NASA).
- Reemplazar nuestro consumo de datos meteorológico (DarkSky dejará de operar en el 2022).
- Búsqueda de información más granular: en lugar de tener datos meteorológicos a nivel mensual, tener esta información cada día.
- Realizar un análisis más exhaustivo para encontrar la ventana de entrenamiento óptima.
- Realizar optimización de hiperparámetros correcta.
- Mejorar la creación de variables adicionales.
- Implementar este proyecto con predicciones o entrenamientos en otros países, como Estados Unidos o Canadá.

- Predicción certera de cambios repentinos en la presencia/ausencia de sequías y evaluaciones más rigurosas