

DESLIZAMIENTOS Y APRENDIZAJE AUTOMÁTICO

José Mauricio Cepeda Rivera¹

Aporte al Seminario de Difusión de Investigaciones UCA 2021.

El trabajo que ahora presento es una obra colectiva, han participado muchos estudiantes de la Universidad de Oslo, colegas del Instituto Geotécnico de Noruega, colegas de la Universidad de Oslo, de la Autoridad de Carreteras y de la Autoridad de Ferrocarriles de Noruega, asimismo, profesionales de la Dirección de Recursos Hídricos y Energía de Noruega.

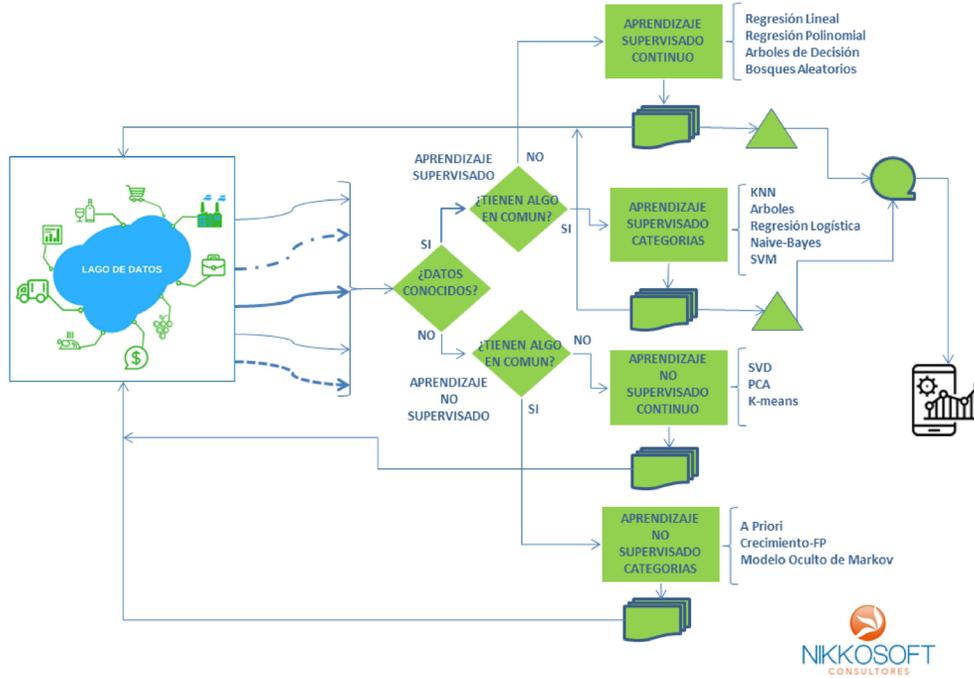
Entraré primero a la parte técnica de la exposición, hablando un poco de la definición de aprendizaje automático o *machine learning* en inglés. En el aprendizaje automático un computador observa datos, luego esos datos sirven para construir un modelo matemático, este modelo se utiliza a la vez como una hipótesis acerca del mundo y una pieza de software para resolver problemas; o sea, predecir algo en base a datos observados.

El aprendizaje automático puede resumirse, por ejemplo, en que basado en lo que una persona ha hecho en internet, sus búsquedas, preferencias de películas, en los correos que ha enviado, las páginas que ha visitado, los libros y otros productos que ha comprado, internet mismo le ofrece productos y servicios adecuados a ello. Lo anterior es una aplicación del aprendizaje automático en redes sociales, pero aquí vamos a verlo aplicado a riesgos de deslizamientos.

En el flujograma del aprendizaje automático, al cual hago referencia solo para efectos ilustrativos, al inicio tenemos un lago de datos, es decir, mucha información sobre usuarios, sus preferencias, sus características; luego, teniendo esos datos podemos tomar dos caminos, el primer caso, que es el más común y sobre él hablaremos en esta exposición, es el aprendizaje supervisado, cuanto tenemos datos conocidos.

¹ Ingeniero Civil (UCA, 1993), obtuvo el grado de Máster en ingeniería civil en el Imperial College de Londres, Reino Unido (1995), en 2009 obtuvo el título de Doctor en Geociencias con la Universidad de Oslo, Noruega. Actualmente, trabaja como Catedrático de geomecánica, amenazas geológicas y deslizamientos para la Universidad de Oslo Noruega.

FLUJOGRAMA APRENDIZAJE AUTOMATICO



Posteriormente, revisamos si esos datos conocidos tienen algo en común y determinamos qué tipo de algoritmo va a ser usado. Un algoritmo simplemente es una serie de pasos que se pueden formular en lenguaje de programación. Podemos tener un aprendizaje supervisado continuo o en categorías; a continuación, podemos escoger diferentes técnicas para llevar a cabo el aprendizaje supervisado. Al final, el objetivo es que con una serie de variables explicativas, vamos a poder predecir una respuesta. Básicamente, el aprendizaje automático nos estaría dando la posibilidad de brindar predicciones a una respuesta, con base en la experiencia de datos conocidos.

Como ya he mencionado, el aprendizaje automático no lo vamos a ver relacionado a medios sociales ni a compras en internet, si no que me concentraré en la problemática de deslizamientos. Para eso, haré algunas definiciones básicas:

Un deslizamiento es un movimiento descendente de suelo y/o rocas debido al efecto gravitacional. El efecto gravitacional está contrapuesto al movimiento debido a erosión, porque en el caso de la erosión, el movimiento por parte de suelo y rocas ocurre como parte de una corriente de agua; mientras que, en el caso gravitacional, el movimiento descendente únicamente ocurre por el peso propio de los materiales sin involucrar una corriente de agua que los arrastre.

Los tipos de materiales en un deslizamiento pueden ser rocas y suelos, estos últimos, desde el punto de vista de la ingeniería, se clasifican en gruesos y finos; por otra parte, los tipos de movimiento o mecanismo, pueden ser caídas o desprendimientos, vuelcos o desplomes, deslizamientos, expansiones laterales, flujos y movimientos complejos.

La clasificación de deslizamiento viene dada por el tipo de movimiento o mecanismo más el tipo de material, es una combinación de estas dos cosas. Por ejemplo, así hablamos de caídas de rocas, deslizamientos de tierra y flujos de detritos, allí está presente el mecanismo y el material.

Se reconocen tres tipos de movimientos/mecanismos: caídas, deslizamientos y flujos. Los deslizamientos pueden ser translacionales y rotacionales; mientras que en los flujos, suele cambiar el estado de los materiales de manera incoherente a su estado original (por ejemplo, cambian de estado sólido a estado líquido o semilíquido), en los cuales, un elemento clave en el mecanismo es la lluvia o el agua en general.

¿Por qué es importante conocer los tipos de deslizamiento? Es importante porque no existe una manera única de estudiar la variedad de deslizamientos mencionados, sino que, los análisis tienen que estar adecuados a cada una de estas variedades.

La experiencia de investigación que voy a mencionar se trata de una alerta temprana contra deslizamientos implementada en Noruega, que fue desarrollada en el curso de varios años. Hago referencia específica a dos publicaciones hechas en 2014 y 2019 (Bell et al., 2014; Devoli et al., 2019), pero en realidad, es un trabajo que se viene realizando desde el año 2009 mediante estudiantes de maestría y doctorado, y a la fecha se contabilizan 12 años de trabajo en el tema.

Como introducción, en Noruega ha habido muchos deslizamientos severos en las últimas décadas, por ejemplo en los años 2000, 2005, 2008, 2011 y 2013. Principalmente deslizamientos de detritos superficiales y flujos de detritos, los cuales son desencadenados por lluvia y derretimiento de nieve. A diferencia de El Salvador, donde el desencadenamiento por lluvia es el más importante o el único importante en cuanto a factores hidrológicos, en el caso de Noruega el desencadenamiento por derretimiento de nieve es significativo, particularmente al principio de la primavera y también en el otoño. Estos deslizamientos han causado daños significativos a redes de transporte (vial y ferroviaria) y edificaciones.

Destaco esto porque yo estudié y trabajé en la UCA por varios años, la considero mi alma mater, de la cual, los pilares de su actividad son la docencia, investigación y proyección social. Aquí quiero entrar en mi exposición con el aspecto de proyección social, porque la investigación en este caso tiene sentido, ya que procura de alguna forma, ayudar a reducir estos daños a futuro.

En Noruega se ha desarrollado un sistema de alerta temprana (SAT) para deslizamientos superficiales y flujos de detritos, a cargo de la Dirección Noruega de Recursos Hidráulicos y Energía, entidad encargada de coordinar el manejo y mitigación de riesgos en este campo. La investigación fue posible porque existe responsabilidad legal, jurídica e institucional de este aspecto a nivel nacional, un organismo facilitador. Por otra parte, todos los esfuerzos de mitigación se han hecho en conjunto con institutos de investigación, consultores y universidades: Universidad de Oslo, Instituto Geotécnico Noruega, entre otros.

Algo que es importante destacar es el financiamiento, el cual viene de los actores interesados. Como mencioné antes, los daños por deslizamientos ocurren en redes de transporte vial y ferroviaria, de modo que la Autoridad de Carreteras y la Autoridad de Ferrocarriles de Noruega jugaron un papel importante, así como entidades del gobierno central a través del Consejo Nacional de Investigación Noruega.

Por otra parte, el sistema de alerta temprana está basado en umbrales hidrometeorológicos, que son una combinación de suministro de agua relativo y el grado de saturación del suelo relativo, presentado como mapas de índices con una resolución de 1 km por 1 km. Es un sistema que está operativo desde octubre de 2013, el cual ha venido mejorando con el tiempo y como resultado de

las investigaciones que se han venido realizando.

El sistema funciona dando alertas diarias a las autoridades de las redes viales y ferroviarias, municipalidades y autoridades de emergencia. Es decir, los usuarios finales del sistema son los interesados en los elementos que potencialmente pueden sufrir daños.

Un problema que ocurría cuando se inició este sistema de alerta en 2013, es que había una sobreestimación de niveles de alerta para algunas áreas que no son susceptibles a deslizamientos, así como subestimación de otras áreas. Esta fue la base de nuestro tema de investigación.

Así, la investigación identificó un problema de investigación, referido a la sobreestimación y subestimación de alertas. El sistema producía falsas alarmas en algunos lugares, mientras que no producía alertas en áreas donde sí ocurrían deslizamientos. Cuando un sistema de alerta produce esos dos errores, significa que debemos mejorar su confiabilidad y funcionalidad. De tal forma que el objetivo de la investigación fue mejorar el sistema de alerta temprana integrando la susceptibilidad a deslizamientos. En ese momento que se realizó el estudio, el elemento susceptibilidad no era parte del sistema.

Las áreas de estudio que se tomaron en cuenta fueron el norte y el sudeste de Noruega, El norte se conforma morfológicamente de montañas, fiordos, valles e islas grandes, donde ocurren numerosos deslizamientos y flujos de detritos; en la base de datos tenemos un total de 759 deslizamientos ocurridos en 284 cuencas hidrográficas.

Por otra parte, el sudeste de Noruega es un área montañosa con valles y extensas tierras bajas, con deslizamientos y flujos de detritos en depósitos cuaternarios en valles. En nuestra base de datos hay registro de 1419 deslizamientos ocurridos en un total de 536 cuencas.

El método utilizado fue un modelamiento del territorio a nivel de cuenca y no por pixel. Esto quiere decir que se toman en cuenta los puntos donde inicia y donde termina el deslizamiento. Un error común era consignar el punto del deslizamiento en el lugar donde finaliza, pero eso no sirve para el análisis ni para mejorar el sistema de alerta temprana. De tal forma, que el análisis es más adecuado cuando se estudia la cuenca en su conjunto y no los pixeles que suelen ser arbitrarios. Con eso se evitan los errores debido a la inexactitud de ubicación del deslizamiento.

El modelo matemático adoptado fueron los modelos aditivos generalizados (GAM). De acuerdo con este planteamiento, se consideran las no linealidades en la relación entre las variables explicativas y las variables de respuesta. Algunas variables pueden omitirse, usarse linealmente o con una función suavizadora. Se estableció una relación de 1:1 de cuencas con deslizamientos y sin deslizamientos. Se utilizaron 50% de cuencas para el entrenamiento y 50% de cuencas para la muestra de prueba.

Por otra parte, la selección de variables estuvo basada en el criterio de información de Akaike (AIC). Es decir, se consideraron como punto de partida una gran cantidad de variables explicativas, pero no todas ellas fueron válidas en el análisis. Finalmente, los modelos se validaron utilizando el criterio AUROC, el cual sirve para establecer mejor las posibilidades de predicción de los modelos.

Las variables explicativas utilizadas fueron:

- Unidades del mapa cuaternario (1:250.000, Servicio Geológico Noruego).

- Tipo de cobertura del terreno (CORINE, 2006, Instituto Noruego Forestal y de Paisajes).
- Lluvia anual promedio (1961-1990, 1 km por 1 km, Instituto Meteorológico Noruego).
- Diversas variables de escorrentía (Base de Datos Nacional de Cuencas, REGINE, NVE).
- Diversos derivados del modelo de elevación digital del terreno, 15 metros por 15 metros (Por ejemplo, pendiente y aspecto) (Instituto Geotécnico Noruego).
- Datos categóricos: proporción calculada de cada clase.
- Datos continuos: media, mediana, desviación estándar.

Por otro lado, la variable de respuesta en este caso era únicamente definir si en las cuencas había habido deslizamientos, es decir, fue una variable binaria.

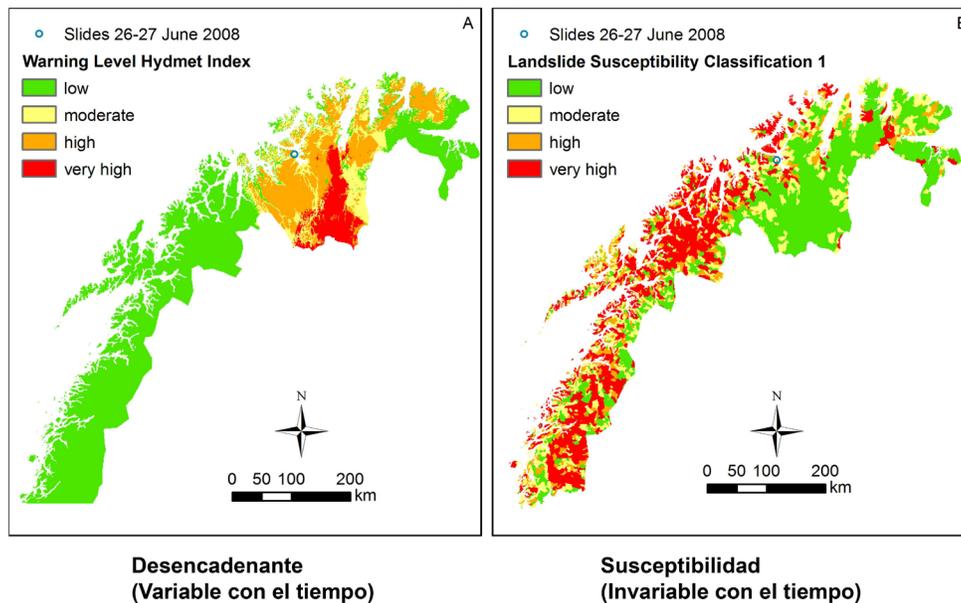
Una vez establecido el modelo, los mapas de susceptibilidad a deslizamientos fueron clasificados en cuatro clases (similar al índice hydmet): susceptibilidad muy alta, alta, moderada o baja. Se hicieron dos versiones de la clasificación.

En cuanto a los resultados para el área norte de Noruega, se obtuvieron altos valores de AUROC de 0.87 y 0.84 en las muestras de entrenamiento y de prueba respectivamente, lo cual indica predictibilidad del modelo. Las variables explicativas finalmente seleccionadas fueron: área de cuenca, pendiente media de cuenca, intervalo de elevación de la cuenca, bosque, espacio abierto con poca o ninguna vegetación, morenas (tipo de material en zonas glaciares), depósito de ríos y arroyos.

Por otra parte, en cuanto al área sudeste de Noruega, también se obtuvieron altos valores de AUROC de 0.88 y 0.86 en las muestras de entrenamiento y de prueba respectivamente. Las variables explicativas que resultaron seleccionadas fueron: área de cuenca, desviación estándar de la pendiente de la cuenca, rango de elevaciones de la cuenca, escorrentía total, lluvia anual promedio, bosque, espacio abierto con poca o ninguna vegetación, áreas agrícolas heterogéneas, depósitos marinos y de fiordos, depósitos fluviales, material intemperizado, turba y marismas.

El aporte de esta investigación fue introducir el elemento de susceptibilidad al sistema de alerta temprana (SAT). La susceptibilidad no varía en el tiempo (mapa de la derecha), porque la susceptibilidad no depende de las condiciones de precipitación instantáneas que pueden llegar a desencadenar un deslizamiento, sino que se refiere a los factores condicionantes para que haya deslizamiento, es decir, condiciones permanentes en una ladera.

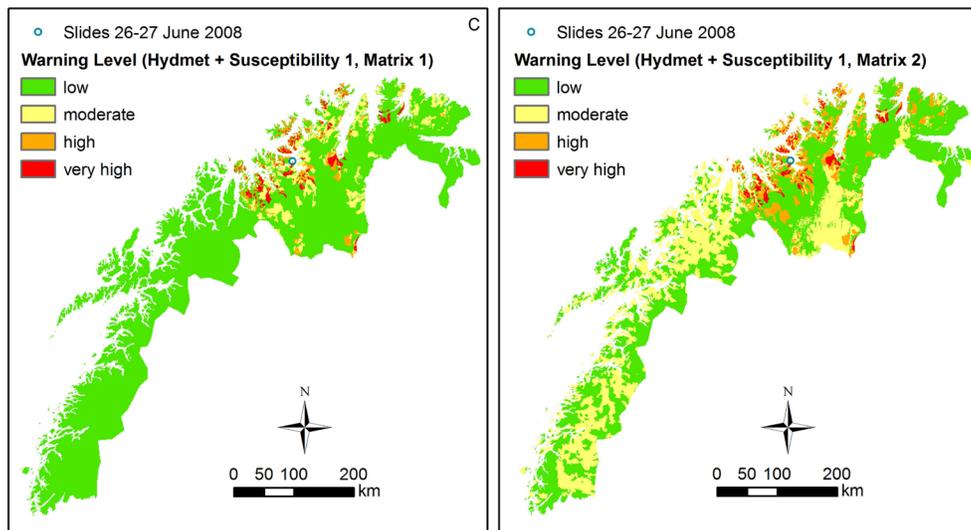
Resultados (SAT)



Mientras que los factores desencadenantes sí varían con el tiempo y se refieren a las condiciones hidrometeorológicas. Antes de la investigación, el sistema de alerta temprana sólo consideraba estos factores desencadenantes (mapa de la izquierda).

Puede verse cómo el mapa de la izquierda muestra niveles de alerta muy altos en zonas donde el nivel de susceptibilidad es muy baja, indicando que en esas zonas se producían seguramente falsas alarmas. Nótese, que también hay zonas en verde donde el nivel de susceptibilidad puede ser alto. Esto indica que deberían darse niveles de alerta diferenciados por cada área específica.

Resultados (SAT)



Como aporte, la investigación estableció un nivel de alerta combinado, mezclando factores de susceptibilidad con factores desencadenantes, para definir finalmente los niveles de alerta en cada región específica, de manera más precisa (segundo grupo de mapas).

Como conclusiones de la investigación, se mencionan los siguientes aspectos:

- El modelamiento de susceptibilidad a deslizamientos en cuencas de primer orden es adecuado para el sistema de alerta temprana por deslizamientos en Noruega.
- La mala exactitud en la ubicación de los puntos de deslizamiento en el inventario (base de datos) no afectó el análisis, dado que los puntos del inventario y el punto real de inclinación se encontraban en la misma cuenca.
- El uso de proporciones de deslizamiento por clase de susceptibilidad funcionó mejor que otros métodos. Además, fue más fácil de entender para las autoridades.
- El nivel de alerta combinado da mejores resultados que solo el índice de desencadenamiento.

Para finalizar esta presentación, quiero dejar planteadas las siguientes preguntas reflexivas sobre los alcances de proyección social en las investigaciones científicas:

- ¿Cuál es la necesidad que se pretende resolver en una investigación?
- ¿Quiénes son los beneficiarios?
- ¿Quiénes son los socios/contrapartes?
- ¿Quién será el dueño del proyecto, producto o servicio?

Estos son los puntos que quería mostrar sobre la investigación, sobre todo relatar la experiencia de esta investigación y cómo ha habido combinación con distintos actores. Muchas gracias por su atención.

Referencias

Bell, Cepeda, Devoli & Glad. (2014). Landslide susceptibility modeling at catchment level for improvement of the landslide early warning system in Norway. Proc. World Landslide Forum 3, Beijing.

Devoli, Bell & Cepeda. (2019). Susceptibility map at catchment level, to be used in landslide forecasting, Norway. Norges vassdrags- og energidirektorat, report No. 1/2019.