



CUARTO CONGRESO INTERNACIONAL
CIENTÍFICO Y TECNOLÓGICO
DE LA PROVINCIA DE BUENOS AIRES

***MACHINE LEARNING APLICADO A
REMOTE SENSING: APLICACIONES EN
GOBERNANZA DIGITAL PARA EL
DESARROLLO SUSTENTABLE***

Instituto de Investigación y Transferencia
en Tecnología (ITT), Escuela de Tecnología,
Universidad Nacional del Noroeste de la Provincia
de Buenos Aires (UNNOBA)

MACHINE LEARNING APLICADO A REMOTE SENSING: APLICACIONES EN GOBERNANZA DIGITAL PARA EL DESARROLLO SUSTENTABLE

J. Tessore, L. Cicerchia, L. Esnaola, H. Ramón, C. Russo, M. Sarobe, S. Serafino, A. Smail*

Instituto de Investigación y Transferencia en Tecnología (ITT), Escuela de Tecnología, Universidad Nacional del
Noroeste de la Provincia de Buenos Aires (UNNOBA)
contacto@itt.unnoba.edu.ar

RESUMEN

El presente trabajo muestra las problemáticas abordadas mediante tecnologías de Machine Learning aplicadas a Remote Sensing, las que pueden servir de soporte a la gobernanza digital para el desarrollo sustentable. Para ello se realiza una revisión bibliográfica de la utilización de estas dos tecnologías con el fin de mostrar cuáles son los avances alcanzados en el área, en qué se está trabajando y cuáles son las aplicaciones relacionadas con la gobernanza electrónica para el desarrollo sustentable.

Palabras clave: *Electronic Governance, Sustainable Development, Machine Learning, Remote Sensing.*

*J. Tessore, L. Cicerchia: becarios de la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC).
H. Ramón, C. Russo: investigadores asociados de la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC).
juanpablo.tessore@itt.unnoba.edu.ar, lucas.cicerchia@itt.unnoba.edu.ar, leonardo.esnaola@itt.unnoba.edu.ar,
hugo.ramon@itt.unnoba.edu.ar, claudia.russo@itt.unnoba.edu.ar, monica.sarone@itt.unnoba.edu.ar,
sandra.serafino@itt.unnoba.edu.ar, ana.smail@itt.unnoba.edu.ar.

INTRODUCCIÓN

Los constantes avances en el área de las tecnologías de la información y la comunicación (TIC) extendieron su alcance a diferentes áreas, entre ellas a los gobiernos. En la actualidad, el mundo se encuentra atravesando una era en la que cada vez más cosas se realizan de manera digital (con diferentes dispositivos y tecnologías); desde actividades culturales, sociales, económicas, hasta actividades políticas. Como consecuencia se producen cambios que en algunos casos magnifican los problemas existentes de división, inequidad, exclusión, fraude, inseguridad, desequilibrio de poder y muchos otros (Janowski, 2015). Así surge el concepto de “gobierno digital” con el fin de “reflejar cómo los gobiernos buscan encontrar soluciones digitales innovadoras a las presiones sociales, económicas, políticas y de otro tipo, y cómo se transforman en el proceso. Comprender y predecir estos cambios es importante para los responsables políticos, ejecutivos gubernamentales, investigadores y todos aquellos que preparan, hacen, implementan o evalúan las decisiones del Gobierno Digital” (Janowski, 2015).

Proveer de información útil para la toma de decisiones en diferentes disciplinas se ha convertido en un paradigma que hoy se encuentra en auge, incluido el gobierno digital. Para eso es necesario realizar un correcto análisis de los datos que las tecnologías proveen y a partir de ello poder obtener información relevante. Las tecnologías generan una gran cantidad de datos que, administrados y procesados debidamente, permiten adquirir mucha información que puede ser útil para la toma de decisiones. La gran cantidad de datos a la que se accede a través de la tecnología requiere de la introducción de técnicas específicas que permitan operar con grandes volúmenes de datos y obtener de ellos información. Así es como se asiste hoy en día a un nuevo concepto o disciplina denominado *Big Data*, el cual hace referencia al manejo de grandes volúmenes de datos, incluyendo su almacenamiento, recolección, búsqueda y análisis, así como también a las técnicas dedicadas a extraer información a partir de los mismos. Para que una máquina pueda extraer información a partir de estos datos, debe comportarse de manera inteligente, es decir, debe ser capaz de resolver problemas del mismo modo que lo hacen los humanos, en base a la experiencia y al conocimiento.

De esto se ocupa la inteligencia artificial (IA), la cual es un área multidisciplinaria, que, a través de las ciencias de la computación, la matemática, la lógica y la filosofía, estudia la creación y diseño de sistemas capaces de resolver problemas cotidianos por sí mismos, utilizando como paradigma la inteligencia humana. Dentro de lo que a la IA concierne, se encuentra el *Machine Learning* (ML), área que engloba un conjunto de técnicas que hacen posible el aprendizaje automático a través del entrenamiento con grandes volúmenes de datos. Es decir, ML trata acerca de computadoras que puedan, mediante diferentes operaciones, aprender y adaptarse en base a la precisión de sus resultados. Esto es traducido en diferentes tipos de algoritmos, los cuales pueden agruparse en tres grandes grupos: *Supervised Learning* o aprendizaje supervisado, *Unsupervised Learning* o aprendizaje no supervisado y *Semi-Supervised Learning* o aprendizaje semisupervisado (Marsland, 2014).

Las técnicas de ML pueden ser aplicadas a todo tipo de datos, los cuales podrían ser los generados por tecnologías como es el caso de *Remote Sensing* o sensadoremoto (RS). Este último configura el campo de la ciencia que estudia y modela los procesos que ocurren en la superficie de la tierra y su interacción con la atmósfera (Kiefe y Lillesand, 2015). Los datos de RS pueden ser adquiridos de diversas formas: vehículos aéreos no tripulados o *Unmanned Aerial Vehicle* (UAV), satélites, etc.; y con diversos sensores como, por ejemplo, el *Landsat Thematic Mapper / Enhanced Thematic Mapper Plus* (TM/ETM+), el espectrorradiómetro de imagen de resolución moderada o *Moderate Resolution Imaging Spectroradiometer* (MODIS), el radar de apertura sintética o *Sintetice Apertura Radar* (SAR), cámaras multi o hiper espectrales, etc. (Yu *et al.*, 2016).

El procesamiento de imágenes de RS posee aplicaciones de la vida real con claros beneficios para la sociedad. Los objetivos principales de RS son (Tuia y Camps-Valls, 2011):

- Monitorear y modelar los procesos sobre la superficie terrestre y su interacción con la atmósfera.
- Medir y estimar las variables geográficas, biológicas y físicas.
- Identificar materiales en la cubierta terrestre y analizar las firmas espectrales adquiridas por sensores satelitales o aéreos.

Todo ello se puede realizar debido a que los materiales en una escena reflejan, absorben y emiten radiación electromagnética de una manera diferente en función de su composición y forma molecular, lo que permite la adquisición de información acerca de una escena (o un objeto específico) a una corta, media o larga distancia (Camps-Valls y Paterna, 2009). De ello resulta una imagen multiespectral que consiste en una adquisición simultánea de imágenes espacialmente coregistradas, en varias bandas contiguas del espectro electromagnético de una plataforma operada remotamente (Kiefe y Lillesand, 2015) (Shaw y Manolakis, 2002). Como se puede observar en D. J. Lary *et al.*, la aplicación de técnicas de ML a RS es relativamente nueva y limitada, abarcando diferentes campos de aplicaciones y ejemplos. Entre otros resultados se pueden obtener mapas de clasificación, predicción, selección y extracción de características del suelo.

Entre las diferentes técnicas de ML que se encuentran, y teniendo en cuenta aquellas más recientemente aplicadas a *Remote Sensing*, se destacan las cinco más utilizadas (Camps-Valls y Paterna, 2009): *Semi-Supervised Learning*, *Manifold Learning* o aprendizaje colector, *Transfer Learning* o aprendizaje de transferencia, *Structured Learning* o aprendizaje estructurado, *Deep Learning* o aprendizaje profundo (DL) y *Active Learning* o aprendizaje activo (AL). Dichas técnicas pueden resultar de gran utilidad y ser el soporte para la toma de decisiones de los gobiernos digitales. A continuación, se desarrollará una breve descripción sobre gobierno digital y cada uno de sus componentes, para luego hacer una revisión bibliográfica sobre qué aplicaciones de ML y RS pueden ser tomadas por el gobierno digital como sustento a la toma de decisiones.

DESARROLLO

Gobernanza electrónica

Según TomaszJanowski, el gobierno digital conlleva, al igual que los cambios en las culturas y las sociedades, una evolución, y esta se desarrolla en cuatro etapas diferentes:

- Digitalización (tecnología en gobierno). Comprende la digitalización de la información gubernamental y automatización de las operaciones y los sistemas de prestación de servicios públicos.
- Transformación (gobierno electrónico). Supone una mejora del funcionamiento interno y la cultura del gobierno y sirve para facilitar la reforma institucional.
- Participación (gobernanza electrónica). Cuyo objetivo es involucrar a los ciudadanos y a otros actores no estatales en la toma de decisiones gubernamentales y en la creación de confianza.
- Contextualización (gobernanza electrónica impulsada por políticas). Implica la habilitación de territorios, comunidades, ciudadanos, etc.; para que puedan llevar a cabo acciones de desarrollo por sí mismos.

El presente trabajo se centrará, en particular, en uno de los aspectos de gobierno digital que es el de “gobernanza electrónica”, más específicamente en el de “gobernanza electrónica para el desarrollo sustentable” o Electronic Governance for Sustainable Development (EGOV4SD). Según E. Estévez *et al.*, EGOV4SD es:

... el uso de TIC para apoyar los servicios públicos, la administración pública y la interacción entre el gobierno y el público, al tiempo que posibilita la participación pública en la toma de decisiones gubernamentales, el desarrollo económico y la protección de los recursos naturales para las generaciones futuras.

El desarrollo sostenible o *Sustainable Development* (SD), tal como lo indica su nombre, promueve el desarrollo teniendo en cuenta el impacto de las iniciativas en los diferentes sectores y su efecto sobre las generaciones futuras.

A nivel mundial cada vez hay mayor consciencia acerca del impacto que tiene el accionar humano tanto en el aumento de la población, como en la disminución del bienestar humano, la pérdida de ecosistemas, el cambio climático global, etc.; lo cual demuestra que los caminos actuales de desarrollo socioeconómico no son sustentables (como se cita en Estévez *et al.*, 2013). Así surge la preocupación de los gobiernos y la necesidad de tomar cartas en el asunto. Reconociendo estas preocupaciones, las Naciones Unidas pidieron que se formulase una “agenda global para el cambio” y la Comisión Mundial de Desarrollo Económico elaboró el informe denominado “Nuestro futuro común” (como se cita en Estévez *et al.*, 2013), que introdujo el concepto de desarrollo sostenible como el desarrollo que satisface las necesidades de la generación actual sin comprometer la capacidad de las generaciones futuras para satisfacer sus propias necesidades.

Así, EGOV4SD como se cita en Estévez *et al.*:

... se aplica para perseguir: mitigación de la pobreza, integración de políticas ambientales, equidad intra e inter generacional, participación pública en la toma de decisiones y otras medidas para superar nuestros límites al crecimiento. Además, se agrupa el SD a través de cuatro dimensiones: 1) social - representa el capital humano y tiene como objetivo fortalecer la coherencia social y la justicia; 2) ambiental - representa el capital natural y tiene como objetivo salvaguardar el medio ambiente; 3) económico - representa el capital humano y tiene como objetivo satisfacer las necesidades materiales, y 4) institucional - representa el capital social y tiene como objetivo la participación y la toma de decisiones colaborativas.

La diferencia clave entre gobierno electrónico y gobernanza electrónica (EGOV) es que el primero se refiere a lo que sucede dentro del gobierno, mientras que el segundo se refiere a todo el sistema involucrado en la gestión de la sociedad (Li *et al.*, 2013). En base a esta definición de EGOV4SD y a los conceptos desarrollados sobre ML y RS, el objetivo del presente trabajo es realizar una revisión bibliográfica de la utilización de *Machine Learning* en *Remote Sensing* con el fin de mostrar cuáles son los avances que hay en el área, en qué se está trabajando y cuáles son las aplicaciones que pueden tener relación con el concepto de EGOV4SD.

En relación con EGOV4SD y las problemáticas que este aborda, existen algunas de ellas que se resuelven mediante estas tecnologías, que podrían ser utilizadas como políticas de gobierno para el desarrollo sostenible.

Aplicaciones de Machine Learning en Remote Sensing

Aquí se desarrolla un análisis de los diferentes estudios que se consideran que cumplen con los objetivos anteriormente mencionados de EGOV4SD y que podrían ser incorporados por los gobiernos dentro de sus políticas. Estos a su vez fueron ordenados por problemáticas.

Detección de focos de pobreza

En este sentido aparecen diferentes estudios que utilizan ML y RS para detectar distintos tipos de problemáticas, como lo es, por ejemplo, la pobreza. La falta de datos fiables en los países en desarrollo y la brecha de generación de datos con respecto a países desarrollados generan un obstáculo importante para el desarrollo sostenible, la seguridad alimentaria y el socorro en casos de desastre (Xie *et al.*, 2015). Allí se propone la utilización de técnica de ML para la extracción de indicadores socioeconómicos a partir de imágenes satelitales crudas. Para ello se utilizan *Transfer Learning* (el cual pertenece al grupo *Semi-Supervised Learning* [Camps-Valls y Paterna, 2009]) y redes neuronales convolucionales o *Convolutional Neural Network* (CNN, el cual pertenece al grupo *Supervised Learning* [Marshall, 2014]) para extraer factores socioeconómicos, aunque el principal desafío, como se mencionó anteriormente, es la obtención de los datos de entrenamiento de los algoritmos. Los autores utilizan las CNN y

Transfer Learning para predecir intensidades de luces nocturnas a partir de imágenes diurnas y realizar un mapeo de la pobreza, identificando entre otras cosas, estructuras artificiales tales como rutas, áreas urbanas y campos sin ninguna supervisión más allá de las luces nocturnas, es decir, sin ningún ejemplo etiquetado de caminos o áreas urbanas. Dichas características son altamente informativas para la cartografía de la pobreza y capaces de acercarse al desempeño predictivo de los datos recolectados en el campo.

N. Jean *et al.* proponen un método alternativo al de realización de encuestas, con el fin de detectar los focos de pobreza en países en desarrollo. Allí, utilizando datos públicamente disponibles, obtienen un “método preciso, barato y escalable para estimar el gasto de consumo y la riqueza de activos a partir de imágenes satelitales de alta resolución”. Este estudio se realizó utilizando *Transfer Learning* sobre imágenes satelitales, donde a partir de entrenar al algoritmo, etiquetando imágenes en unas mil categorías, lograron estimar la actividad económica, detectando intensidades de luz nocturna en dichas imágenes. El modelo se aplicó sobre cuatro países del África: Nigeria, Tanzania, Uganda, Malawi y Ruanda; y para validar el modelo se utilizaron estudios de dominio público, donde para la medición de los gastos de los hogares se empleó el estudio de mediciones del nivel de vida o *Living Standards Measurement Study* (LSMS) del Banco Mundial y para medir la riqueza se utilizaron los índices de activos, extraídos de los estudios demográficos y de salud o *Demographic and Health Surveys* (DHS), los cuales son la fuente primaria de estadísticas de salud en muchos de los países en desarrollo. Este método de ML muestra cómo un algoritmo puede ser entrenado para identificar características de una imagen y puede llegar a explicar hasta el 75 % de la variación en los niveles económicos del área geográfica que se está estudiando.

Análisis, control y predicciones sobre los cultivos

En la actualidad, las limitaciones que imponen el cambio climático junto con el creciente aumento de la demanda mundial de alimentos hacen que muchas investigaciones busquen dar soluciones a este tipo de problemáticas y que estas soluciones sirvan de soporte para la toma de decisiones por parte de todos los actores responsables en el tema, incluidos los gobiernos, y en un sentido más amplio que contribuyan a un uso responsable del medio ambiente.

Los productos derivados del agro requieren de una gestión adecuada de la producción y de los factores involucrados en ella, donde la agricultura de precisión (AP) es la respuesta más adecuada para lograr dichos requerimientos de producción (Tilman *et al.*, 2011). La AP se trata de la administración del campo basado en la observación, la medición y la respuesta frente a la variabilidad, temporal y espacial, inter o intra cultivo (Wang *et al.*, 2015). La AP propone un manejo de la tierra y el cultivo de forma selectiva de acuerdo con las necesidades de este, junto con una correcta administración de las semillas y los agroquímicos con el fin de aumentar la producción, mejorar los costos y hacer un uso responsable del medio ambiente (Fernández-Quintanilla, 2002), (Blackmore, 2002), (García y Flego, 2008). Por ejemplo, predecir el rendimiento de los cultivos antes de su cosecha ayudaría a los políticos del sector y a los agricultores a tomar medidas apropiadas en su comercialización y su

almacenamiento, ayudando además a las industrias asociadas en la planificación de la logística. Así es que M. Stas *et al.* describen la aplicación de dos métodos de ML para modelar el rendimiento de trigo de invierno en la provincia de Henan, en China.

Gracias a que se ha demostrado que la vegetación posee una firma espectral, es decir que refleja diferentes espectros electromagnéticos, es que se ha logrado determinar que estas firmas son buenos estimadores de la biomasa de los cultivos, el rendimiento, la copa de las plantas, el índice de área foliar (IAF), el contenido de clorofila y el envejecimiento de las plantas. En base a ello, como mencionan A. González-Sánchez *et al.* y S. Veenadhariet *al.*, es que podría llevarse a cabo una supervisión del suelo, el cultivo y el clima en un campo, y proporcionar un sistema de apoyo a las decisiones, que sea capaz de aprender, entre otras cosas, a predecir el rendimiento de los cultivos aplicando técnicas de ML como *Regression Trees*, *Artificial Neural Networks* (Marsland, 2014) o *Support Vector Regression* (las cuales pertenecen al grupo *Semi-Supervised Learning*). Asimismo, como indica S. Dimitriadiset *al.*, en base a un mapeo del estado del cultivo se pueden determinar tratamientos específicos como puede ser el riego, la aplicación de fertilizantes o pesticidas para diferentes partes del campo, lo que permite hacer un uso responsable de los recursos y del medio ambiente. Por último, otro trabajo usa la técnica de ML llamado *Random Forest* (el cual pertenece al grupo *Supervised Learning* [Marsland, 2014]) para extraer datos fenotípicos del cultivo de soja, obteniendo características geométricas de los cultivos, las cuales poseen altas correlaciones con el rendimiento final del cultivo. Así lograron mejorar la precisión de la estimación de rendimiento y la eficiencia de registro de madurez del cultivo, con respecto a los métodos tradicionales de medición (Yuet *al.*, 2016).

Análisis, control y estudios sobre el suelo

Uno de los aspectos que más se han estudiado en cuanto a RS es la cobertura de la tierra, ya que esta ha sido reconocida como una variable fundamental en varios estudios a escala mundial para el medio ambiente y el cambio climático. El suelo posee diferentes propiedades, las cuales proporcionan servicios ecosistémicos: filtra agua, suministra nutrientes a las plantas, nos provee alimentos, fibra y energía, apoya la salud humana y la vivienda, almacena el carbono, regula las emisiones de gases de efecto invernadero y afecta al clima. Los avances en la degradación del suelo y la urbanización amenazan los equilibrios agroecológicos y la seguridad alimentaria, y es por ello que es importante saber más sobre el suelo, para manejarlo y preservarlo de manera sostenible para las generaciones futuras. Para ello R. A. ViscarraRosset *al.* proponen diferentes métodos de ML, por ejemplo para hacer mapeos, monitoreo y modelado de suelo. Otro trabajo relacionado es el de J. J. Maynard y M. R. Levi, donde proponen un método de selección de imágenes hipertemporales: tomar cientos de imágenes en el tiempo debido a que estas dependen del tiempo (clima) del día en que fueron tomadas, y a partir de las imágenes y los datos del clima tener una "huella digital" espectral que contenga la relación suelo-vegetación. Esta huella digital es importante debido a que se encuentra directamente relacionada con una gama de propiedades del suelo. El estudio se

realizó sobre un paisaje semiárido del sudeste de Arizona en Estados Unidos, donde se utilizaron unas imágenes de unos 28 años, y aplicando el algoritmo de ML llamado *Support Vector Machine* (el cual pertenece al grupo *Supervised Learning* [Marsland, 2014]), lograron obtener un buen predictor de textura del suelo. Asimismo, L. Yuet *al.* propone extraer diferentes características de distintos sensores usando dos técnicas de ML, la ya mencionada *RandomForest* y K vecinos cercanos o *K-Nearest Neighbours* (que pertenece al grupo *Semi-Supervised Learning* [Marsland, 2014]). Por último, podemos destacar el trabajo de B. Heunget *al.*, donde realiza un estudio sobre el suelo analizando diferentes técnicas de ML aplicadas a RS, allí se pueden ver diez algoritmos de clasificación para la predicción de unidades taxonómicas del suelo (clasificación del suelo en base a diferentes características) en el Valle del Bajo Fraser, Columbia Británica, Canadá. Allí aplica una variedad de técnicas de ML como *Regression Trees*, *Random Forest*, *K-Nearest Neighbours*, *Artificial Neural Network*, *Support Vector Machine*, entre otros.

Otro tipo de tratamiento de suelo es el caso del estudio de derrumbes que realizaron en la región de Yangtze Three Gorges en China (Li *et al.*, 2015). Este surgió a partir de más de 2500 localidades en laderas inestables, que a menudo sufren de deslizamientos de la tierra, y de la construcción de una represa en dicha área que aumentó la probabilidad de desastres geológicos. El estudio allí realizado logra brindar información absolutamente necesaria para ayudar a evaluar y cuantificar el riesgo de deslizamientos. Para ello han realizado diferentes pruebas en las que compararon análisis basados en píxeles de los datos de la tecnología *Light Detection and Ranging* (LIDAR) (US Department of Commerce) junto con análisis de imágenes basadas en objetos, y luego cómo estos responden a dos algoritmos de ML diferentes: *Random Forest* y *Support Vector Machine*. Otro trabajo que también busca cartografiar, pero además analizar áreas sujetas a peligros naturales como el caso anterior en la región de Yangtze Three Gorges en China, es el de J. Smeckaert *et al.* El cual realiza modelado digital de Terreno con el fin de realizar un modelado del riesgo de inundación. Allí, al igual que en el anterior trabajo mencionado, utiliza la tecnología LIDAR para dicho modelado, ya que esta proporciona una nube de puntos 3D que facilitan la reconstrucción, y a partir de la utilización del método *Support Vector Machine* logra discriminar las áreas de tierra de las de agua. Logra así resultados de una clasificación con alta precisión, superior al 95 % en la mayoría de las áreas costeras y más de un 90 % para los ríos.

Existen además trabajos que permiten guiar la toma de decisiones con respecto a la exploración mineral, la agricultura y la salud ambiental, a través de la realización de mapas o químicos, como es el caso del trabajo de C. Kirkwood *et al.* Allí presentan un trabajo como alternativa al análisis químico del suelo, ya que el método tradicional resulta demasiado costoso. Dicho mapa lo realizan sobre suelos del sudoeste de Inglaterra y para ello utilizan *Regression Forest* (variante de *Random Forest*, que pertenece al grupo *Supervised Learning*), junto con RS de alta resolución y datos de estudios geofísicos.

Relacionado con el cambio climático, surge un trabajo de Neal J. Pasticket *al.* Sobre regiones de latitudes altas, que experimentan rápidos y extensos cambios en la composición y función de los ecosistemas, como resultado del aumento de la temperatura

promedio del aire. Este aumento de temperatura lleva a un descongelamiento y degradación del permafrost (capa del suelo permanentemente congelada en las regiones polares), y de manera indirecta afecta a los ecosistemas, la socioeconomía y al ciclo del carbono (paso cíclico del carbono del aire a las plantas que lo absorben, los animales que ingieren las plantas, y al aire nuevamente por la descomposición de los animales). Se utilizaron los algoritmos de ML llamados *Decision y Regression Trees* (los cuales pertenecen al grupo *Semi-Supervised Learning* [Marsland, 2014]) para cuantificar los cambios en la distribución del permafrost en diferentes escenarios climáticos futuros suponiendo que no hay otros cambios en los factores biofísicos. Allí los modelos indicaron que el permafrost cerca de la superficie subyace al 38 % de Alaska continental y que el permafrost cercano a la superficie desaparecerá del 16 al 24 % del paisaje para fines del siglo XXI.

Análisis, control y monitoreo de plantaciones

En este aspecto RS ofrece oportunidades para recopilar e interpretar automáticamente grandes cantidades de datos legibles por sistemas informáticos que son útiles para la conservación de la naturaleza y el monitoreo de la biodiversidad. Las encuestas que se realizan de manera manual en la gestión de la conservación de la naturaleza son costosas, requieren mucho tiempo y pueden ser complementadas y simplificadas mediante el uso de RS (Moran *et al.*, 2017). Así pueden realizarse estudios de plantaciones como el que proponen E. B. Görgens *et al.*, que evalúa el desempeño de tres herramientas de ML para predecir el volumen de las plantaciones forestales de rápido crecimiento, basadas en métricas estadísticas de vegetación extraídas de un estudio de escáner láser aerotransportado o *Airborne Laser Scanning* (ALS), el cual está basado en la tecnología LIDAR. El trabajo fue realizado sobre un bosque de plantación comercial que posee dos variedades de eucalipto y que se encuentra en un área del estado de San Pablo. Este abarca aproximadamente 1138 hectáreas y está caracterizado por una topografía montañosa que oscila entre 589 y 1294 metros sobre el nivel del mar. La predicción se realizó mediante la utilización de redes neuronales o *Neural Networks*, *Random Forest* y *Support Vector Regression*, y, como se mencionó, se utilizaron datos de *Airborne Laser Scanning*. Otro trabajo interesante es el de M. Singh *et al.* sobre bosques tropicales, que, si bien estos cubren solo el 6 % de la superficie de la Tierra, constituyen un reservorio crucial de biodiversidad. Los bosques, a su vez, proporcionan una multitud de servicios ecosistémicos como el almacenamiento de carbono y el suministro de productos forestales a las comunidades locales –por ejemplo, productos derivados de la madera–. La pérdida de bosques y las perturbaciones antropogénicas asociadas contribuyen aproximadamente entre el 6 % y el 17 % de las emisiones globales de carbono (como se cita en Singh *et al.*, 2015). En base a todos estos datos es que este trabajo realiza la cuantificación de la biomasa aérea o Above Ground Biomass (AGB) almacenada en los bosques, ya que a partir de ella se puede estimar el almacenamiento de carbono de un ecosistema dado, a partir de la utilización de imágenes de alta resolución que provee Google Earth e imágenes aéreas a las cuales se le aplicaron dos

algoritmos de ML (con el fin de determinar cuál obtenía mejores resultados) *Random Forest* y *Support Vector Regression*.

RESULTADOS

Sumado al sensado remoto, se ha trabajado con tecnologías de relevamiento de datos a campo de forma automatizada. Incorporando el uso de una plataforma integrada de navegación y sensado, el uso de imágenes digitales y video, y la utilización de técnicas de visión artificial y procesamiento de señales, se pueden obtener datos biofísicos de un cultivo durante su desarrollo evolutivo, que sirvan luego para su monitoreo, análisis y evaluación, permitiendo mejorar el estudio de aspectos cualitativos y cuantitativos en relación a su variabilidad fenológica, morfológica, fisiológica, temporal y espacial.

La utilización de técnicas de procesamiento digital, tanto en sensado remoto como a campo, permiten aportar un valor agregado a las tecnologías ya existentes en esta área. En ambos casos existen diferentes tipos de procesamientos de las imágenes con el fin de calcular índices de vegetación para los diferentes tipos de cultivos (a partir de su firma espectral) con los que se desee trabajar, y así obtener información sobre el cultivo que se está estudiando.

En cuanto a sensado a campo se pueden capturar imágenes que permitan la obtención de datos, el monitoreo y la evaluación, así como capturar video que se procese en línea (mediante la utilización de un algoritmo de *tracking* invertido) para la detección de los surcos (líneas de plantación) (Álvarez *et al.*, 2017). Dichos datos pueden ser capturados por diferentes sensores, por ejemplo, cámaras de luz visible para la obtención de las imágenes que luego serán procesadas para el monitoreo y análisis de parámetros fenológicos. Y también una cámara multispectral, la cual permite obtener una imagen con tres bandas de frecuencia diferentes (*Red*, *Green* e *IRC*) que permite calcular diversos índices de vegetación.

A partir de estos mecanismos de sensado, se ha trabajado en cultivos con variabilidad genotípica con el fin de obtener mejoras genéticas del cultivo (en nuestro trabajo de maíz)(Russo *et al.*, 2016), que permitan que estos tengan mejores respuestas a los diferentes factores ambientales, lo que llevaría a una mejor respuesta del cultivo y en consecuencia a una reducción, no solo de los costos de la utilización de productos agregados, como pueden ser el uso de fertilizantes, pesticidas o la utilización de riego artificial, sino también a un menor impacto en el ambiente. A partir de ello, se han desarrollado algoritmos para *stand* de plantas (detección y conteo de plantas) (Russo *et al.*, 2016). Además, en relación con información agronómica de los ensayos (conjunto de pares de líneas de cultivos que poseen diferentes variaciones genéticas), se han desarrollado los algoritmos para la determinación del porcentaje de emergencia, es decir en base a cuantas semillas se plantaron, determinar cuántas de ellas llegaron a germinar, donde se evalúan los mismos aspectos que en el punto anterior. Allí, en cuanto al procesamiento, se han utilizado técnicas basadas en conversión a esquemas de color HSI, umbralización basada en color, morfología matemática binaria, segmentación basada en contornos; y la etapa de análisis se basa en técnicas de identificación de características y

reconocimiento de patrones. Otra de las problemáticas abordadas con esta tecnología ha sido el relevamiento del vigor inicial, es decir la medición de biomasa (tamaño) en los estadios iniciales de implantación del cultivo (cuando el cultivo desarrolló su primera y segunda hoja) determinando en forma comparativa entre las diferentes parcelas que poseen diferentes variaciones genéticas. Por último, el cálculo del índice de verdor diferencial normalizado (NDVI) mediante la utilización de la cámara multiespectral, donde a partir de la imagen mediante el uso de las bandas del rojo y el infrarrojo cercano del espectro electromagnético.

Todos estos aspectos, entre otros, permiten evaluar, cómo es la competencia del cultivo por los nutrientes, cómo influyen los factores ambientales como temperatura, humedad ambiente, humedad del suelo, etc., tanto para su germinación inicial (cuando desarrolla sus primeras hojas) como a la hora del desarrollo y evolución a lo largo de su vida, que, como se mencionó anteriormente permiten *a posteriori*, tener un menor impacto económico y ambiental.

CONCLUSIÓN

Con el correr de los años, los gobiernos han tomado cada vez mayor consciencia de que debían incorporar dentro de sus políticas el uso de las TIC, y que estas a su vez ayudarían a resolver diferentes problemáticas, muchas veces de un modo más económico y eficiente, o hasta incluso solucionar aquellas que de otro modo no se podrían resolver o bien resultaría demasiado costoso. A ello se le suman conceptos como el de EGOV4SD que no solo implica el uso de las TIC como política de gobierno sino que también tiene como objetivo el desarrollo sustentable.

En este trabajo se expusieron dos tecnologías diferentes con un amplio campo aún por explorar. Por un lado, ML, cuyo conjunto de técnicas permiten enfrentar distintas problemáticas de diferentes formas y, por otro lado, RS, que permite recopilar e interpretar grandes áreas y hacerlas legibles mediante la informática. Asimismo, se pudo observar que la combinación de ambas permite obtener muchísima información, que puede ser útil para la toma de decisiones de los gobiernos, pudiendo colaborar además en gran medida, con el desarrollo sustentable.

Ambas tecnologías combinadas permiten, como se mencionaba anteriormente, resolver muchas problemáticas de diferentes modos y en combinación con otras tecnologías. Solo se espera que de a poco dichas tecnologías sean incorporadas por los gobiernos y que puedan aplicarlas para sí, tal como se ha visto en los *papers* citados, configurando de este modo un avance en relación con el campo que aún dichas tecnologías tienen por explorar y descubrir.

Por todo lo expuesto es que actualmente dentro del instituto se desarrollan las líneas de investigación y transferencia abordadas anteriormente, enmarcadas en los dos proyectos acreditados con que cuenta el ITT.

BIBLIOGRAFÍA

- ÁLVARIZ, E.; SERAFINO, S.; CICERCHIA, B.; BALMER, A.; RUSSO, C. y RAMÓN, H. (2017). "Algoritmo de *tracking* invertido para el relevamiento de cultivos a campo mediante visión artificial y robótica". Instituto en Investigación y Transferencia en Tecnología. Universidad Nacional del Noroeste de la Provincia de Buenos Aires. Junín, Buenos Aires, Argentina.
- BLACKMORE, S. (enero, 2002). "Developing The Principles of Precision farming". The Centre of Precision Farming, Dinamarca.
- CAMPS-VALLS, G. (2009). "Machine Learning in Remote Sensing Data processing". *Mach. Learn. Signal Process. - IEEE International Workshop*, pp. 1-6.
- DIMITRIADIS, S. y GOUMOPOULOS, C. (2008). "Applying Machine Learning to Extract New Knowledge in Precision Agriculture Applications". *Proc. - 12th Pan-Hellenic Conf. Informatics*, pp. 100-104.
- ESTÉVEZ, E.; JANOWSKI, T. y DZHUSUPOVA, Z. (2013). "Electronic Governance for Sustainable Development: How EGOV Solutions Contribute to SD Goals". *Proc. 14^a Annu. Int. Conf. Digit. Gov. Res.*, pp. 92-101.
- FERNANDEZ-QUINTANILLA, C. (2002). "Agricultura de precisión". *Ciencia y Medio Ambiente*, Segundas jornadas científicas sobre medio Ambiente del CCMA-CSIC, pp. 187-194.
- GARCÍA, E. y FLEGO, F. (2008). "Tecnología agropecuaria. Agricultura de precisión". *Revista Ciencia y Tecnología*, n.º 8. Recuperado de <<http://www.palermo.edu/ingenieria/downloads/pdfwebc&T8/8CyT12.pdf>>
- GONZÁLEZ-SÁNCHEZ, A.; FRAUSTO-SOLÍS, J. y OJEDA-BUSTAMANTE, W. (2014). "Predictive Ability of Machine Learning Methods for Massive Crop Yield Prediction". *Spanish Journal of Agricultural Research*, vol. 12, n.º 2, pp. 313-328.
- GÖRGENS, E. B.; MONTAGHI, A. y RODRIGUEZ, L. C. E. (2015). "A Performance Comparison of Machine Learning Methods to Estimate the Fast-Growing Forest Plantation Yield Based on Laser Scanning Metrics". *Computers and Electronics in Agriculture*, vol. 116, pp. 221-227.
- HEUNG, B.; HO, H. C.; ZHANG, J.; KNUDBY, A.; BULMER, C. E. y SCHMIDT, M. G. (2016). "An Overview and Comparison of Machine-Learning Techniques for Classification Purposes in Digital Soil Mapping". *Geoderma*, vol. 265, pp. 62-77.
- JANOWSKI, T. (2015). "Digital Government Evolution: From Transformation to Contextualization". *Government Information Quarterly*, vol. 32, n.º 3, pp. 221-236.
- JEAN, N.; BURKE, M.; XIE, M.; DAVIS, W. M.; LOBELL, D. B. y ERMON, S. (2016). "Combining satellite imagery and machine learning to predict poverty". *Science*, vol. 353, n.º 6301, pp. 790-794.

- KIEFE, T. M. y LILLESAND, R. W. (2015). *Remote Sensing and Image Interpretation* (7.^a ed.). Nueva Jersey, EE. UU.: Wiley International Editions.
- KIRKWOOD, C.; CAVE, M.; BEAMISH, D.; GREBBY, S. y FERREIRA, A. (2016). "A Machine Learning Approach to Geochemical Mapping". *J. Geochemical Exploration*, vol. 167, pp. 49-61.
- LARY, D. J.; ALAVI, A. H.; GANDOMI, A. H.; y WALKER, A. L. (2016). "Machine Learning in Geosciences and Remote Sensing". *Geoscience Frontiers*, vol. 7, n.º 1, pp. 3-10.
- LI, S.; WANG, W. C.; LI, L. W. y ZHOU, J. J. (2013). "Survey of Support Vector Machine in the Processing of Remote Sensing Image". *Advanced Materials Research*, vol. 774-776, pp. 1567-1572.
- LI, X.; CHENG, X.; CHEN, W.; CHEN, G. y LIU, S. (2015). "Identification of Forested Landslides Using LiDarData, Object-based Image Analysis, and Machine Learning Algorithms". *Remote Sensing*, vol. 7, n.º 8, pp. 9705-9726.
- MARSLAND, S. (2014). *Machine Learning: An Algorithmic Perspective*(2.^a ed.). Boca Raton, Florida, EE. UU.: CRC Press Taylor & Francis Group.
- MAYNARD, J. J. y LEVI, M. R. (2017). "Hyper-Temporal Remote Sensing for Digital Soil Mapping: Characterizing Soil-Vegetation Response to Climatic Variability". *Geoderma*, vol. 285, p. 94-109.
- MORAN, N.;NIELAND, S.;TINTRUP GEN. SUNTRUP, G. Y KLEINSCHMIT, B. (2017). "Combining Machine Learning and Ontological Data Handling for Multi-Source Classification of Nature Conservation Areas". *International Journal of Applied Earth Observationand Geoinformation*, vol. 54, pp. 124-133.
- PASTICK, N. J.; JORGENSON, M. T.; WYLIE, B. K.; NIELD, S. J.; JOHNSON, K. D. y FINLEY, A. O. (2015). "Distribution of Near-surface Permafrost in Alaska: Estimates of Present and Future Conditions". *Remote Sensing of Environment*, vol. 168, pp. 301-315.
- RUSSO, C.; RAMÓN, H.; ALONSO, N.; CICERCHIA, B.; ESNAOLA, L. y TESSORE, J. (abril, 2016). "Tratamiento masivo de datos utilizando técnicas de *Machine Learning*". *XVIII Workshop de Investigadores en Ciencias de la Computación (WICC 2016)*, Entre Ríos, Argentina, pp. 131-134.
- SHAW, G. y MANOLAKIS, D. (2002). "Signal Processing for HyperspectralImage Exploitation". *IEEE Signal Processing Magazine*, vol. 19, n.º 1, pp. 12-16.
- SINGH,M.;EVANS, D.;FRIESS,D. A.;TAN, B. S. y NIN, C. S. (2015). "Mapping Above-Ground Biomass in a Tropical Forest in Cambodia Using Canopy Textures Derived from Google Earth".*Remote Sensing*, vol. 7, n.º 5, pp. 5057-5076.
- SMEECKAERT, J.; MALLET, C.; DAVID, N.; CHEHATA, N. y FERRAZ, A.(2013). "Large-Scale Classification of Water Areas Using Airborne Topographic LidarData" *Remote Sensing of Environment*, vol. 138, pp. 134-148.

- STAS, M.; VAN ORSHOVEN, J.; DONG, Q.; HEREMANS, S.; BRUSSELS, I. y ZHANG, B. B. (2016). "A Comparison of Machine Learning Algorithms for Regional Wheat Yield Prediction Using NDVI Time Series of SPOT-VGT".
- TILMAN, D.; BALZER, C.; HILL, J. y BEFORT, B. L. (2011). "Global Food Demand and the Sustainable Intensification of Agriculture" *Proc. Natl. Acad. Sci. United States Am.*, vol. 108, n.º 50, pp. 20260-20264.
- TUIA, D. y CAMPS-VALLS, G. (2011). "Recent Advances in Remote Sensing Image Processing". *16ª IEEE Int. Conf. (2009). Image Processing (ICIP)*, pp. 3705-3708.
- US DEPARTMENT OF COMMERCE, "What is LIDAR?". Recuperado de <<http://oceanservice.noaa.gov/facts/lidar.html>>
- VEENADHARI, S.; MISRA, B. y SINGH, C. D. (2014). "Machine Learning Approach for Forecasting Crop Yield Based on Climatic Parameters". *International Conference on Computer Communication and Informatics: Ushering Technologies of Tomorrow, Today*, pp. 1-5.
- VISCARRA ROSSEL, R. A.; BEHRENS, T.; BEN-DOR, E.; BROWN, D. J.; DEMATTÉ, J. A.; SHEPHERD, K. D.; Ji, W. (2016). "A Global Spectral Library to Characterize the World's Soil". *Earth-Science Reviews*, vol. 155, pp. 198–230.
- WANG, Y.; LEE, K.; CUI, S.; RISCH, E. y LIAN, J. (2015) "Agriculture Robot and Applications". En *Future Information Engineering and Manufacturing Science: Proceedings of the 2014 International Conference on Future Information Engineering and Manufacturing Science (FIEMS 2014)*, pp. 43-46.
- XIE, M.; JEAN, N.; BURKE, M.; LOBELL, D. y ERMON, S. (2015). "Transfer Learning from Deep Features for Remote Sensing and Poverty Mapping". *30th AAAI Conference on Artificial Intelligence*, p. 16.
- YU, L.; FU, H.; WU, B.; CLINTON, N. y GONG, P. (2016) "Exploring the Potential Role of Feature Selection in Global Land-Cover Mapping". *International Journal of Remote Sensing*, vol. 37, n.º 23, pp. 5491-5504.
- YU, N.; LI, L.; SCHMITZ, N.; TIAN, L. F.; GREENBERG, J. A. y DIERS, B. W. (2016). "Development of Methods to Improve Soybean Yield Estimation and Predict Plant Maturity with an Unmanned Aerial Vehicle Based Platform". *Remote Sensing of Environment*, vol. 187, pp. 91-101.