

La academia en la transformación IA de la empresa

Instituto de Computación, UDELAR





INSTITUTO

El InCo y su relación con la
industria

01

FUNDAMENTOS

Ciencia de datos, Inteligencia
artificial

02

AGENDA

03

IMPLANTACIÓN EN LA EMPRESA

Estrategía, roles y recursos

04

CASOS DE ESTUDIO

Ejemplos en Telcos y otras
industrias





Pablo Rodríguez Bocca

- **Investigador, Científico**
 - **Modelado Predictivo**
 - **Optimización e investigación operativa**
 - **Análisis de datos en redes**
- PhD en Computación (France+Uruguay). Msc (Uruguay). Ingeniero Eléctrico y en Computación (Uruguay)
- Profesor agregado (grado 4, InCo, UDELAR, Uruguay)
- Asesor en ANTEL (mayor empresa de comunicaciones de Uruguay)
- Co-fundador de GoalBit (video streaming) y Agronóstico (agritech)



01

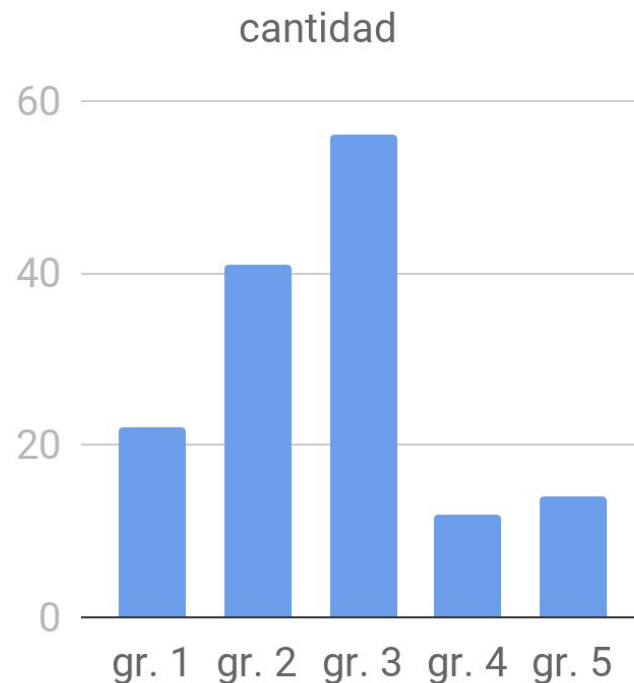
EL INSTITUTO

El Instituto de Computación y su relación con la Industria



Docentes

- **145 docentes**
 - **dedicación total**
 - 41 (28%)
 - **posgrado**
 - 37 doctores, 38 magisters, 4 especialistas (54%)
 - **16 equipos de trabajo**





Posgrado y educación permanente

1/2

- **especializaciones y maestrías profesionales** en Ingeniería de software, seguridad informática, sistemas de información
- **posgrados académicos** en informática (y bioinformática)
- participación en otros posgrados
- **cursos de posgrado (33 cursos) y de actualización profesional (26 cursos)**
- **163 estudiantes de posgrado**

- nuevas iniciativas (en trámite)
 - **especialización en Ciencia de Datos**
 - **maestría en Ciencia de Datos Aplicada**
 - **maestría en Ciencia de Datos y Aprendizaje Automático**
 - diploma en Optimización
 - posgrados para Profesores en Informática (egresados INET)



Investigación y Extensión

1/2

- **16 equipos de investigación**
 - la mayoría con conocimiento/aplicaciones de DS-AI
- **41 proyectos de investigación**
- **vinculación y asesoría** al Estado y empresas
 - 10 convenios y asesoramientos
 - extensión: 4 proyectos, decenas de jornadas, talleres y charlas
- **100+ publicaciones** por año en revistas y conferencias
- **163 estudiantes** de posgrado dirigidos



Investigación y Extensión: equipos

2/2

1. Algoritmos, Análisis de Algoritmos y Combinatoria Analítica
- 2. Centro de Cálculo**
- 3. Concepción de Sistemas de Información**
4. COAL
5. Didáctica de la Informática
6. Grupo de Computación Gráfica
7. Grupo de Ingeniería de Software
8. Grupo de Seguridad Informática
9. Heterogeneous Computing Laboratory
- 10. Investigación Operativa**
11. Laboratorio de Integración de Sistemas
- 12. MINA - Network Management | Artificial Intelligence**
13. Métodos Formales y Programación Funcional
14. Núcleo de Teoría de la Información
- 15. Procesamiento de Lenguaje Natural**
- 16. Sistemas de Información Semánticos**

02

FUNDAMENTOS

Ciencia de datos, Inteligencia artificial





Ciencia de Datos *(data science - DS)*

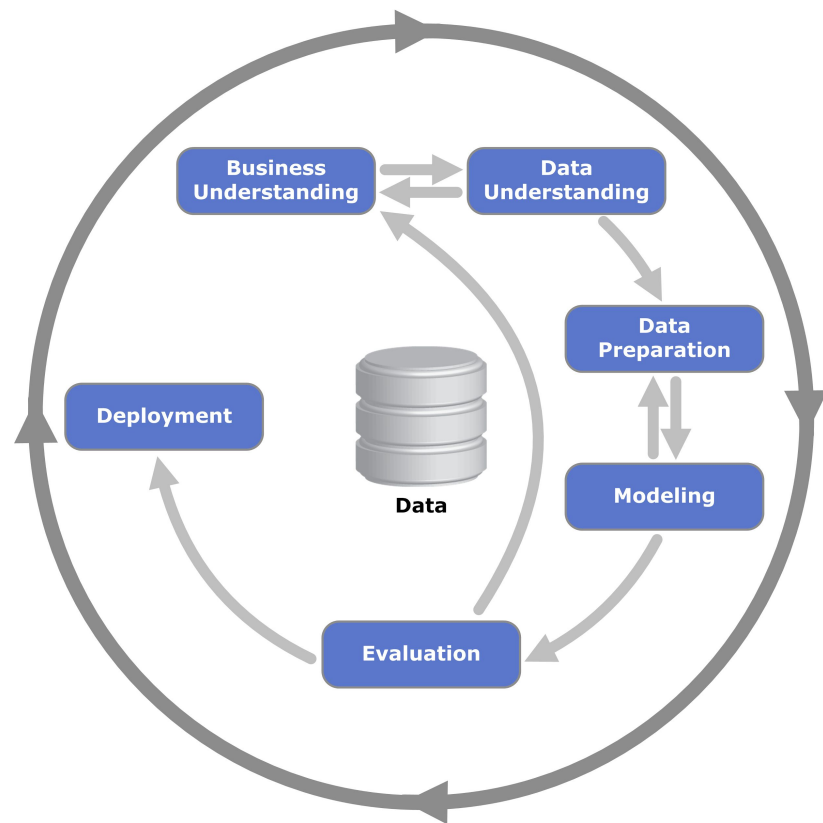
Disciplina que busca **extraer conocimiento** de forma sistemática y computacionalmente eficiente a partir de los **datos**

Fuente: Maestría en Ciencia de Datos y Aprendizaje Automático, Fing, UDELAR



Método científico aplicado a datos

1. preguntarse
2. formular hipótesis
3. experimentar
4. evaluar y mejorar
5. comunicar



Tipos de Análisis y Resultados

1. Descriptivo

- explorar, entender
- *documentos*

2. Predictivo

- futuro y datos faltantes
- automatizar
- *documentos + software*

3. Prescriptivo

- optimizar
- *software*





Inteligencia Artificial

Inteligencia Artificial General : imitar todas las capacidades de los humanos

Inteligencia Artificial Estrecha : aplicado a una sola tarea

Fuente: -





Aprendizaje Automático *(machine learning - ML)*

Disciplina que da a las **computadoras** la **habilidad de aprender** sin ser programados explícitamente

Fuente: Arthur Samuel, 1959



Aprendizaje supervisado

- aprender una función de mapeo

entrada (X)	→ salida (Y)	aplicación
email	¿spam?	filtro spam
audio	transcripción a texto	reconocimiento de voz
Inglés	Español	traducción
imagen de un teléfono	¿defecto?	inspección visual
imagen, radar, gps	posición de otros autos	manejo autónomo
banner, dato usuario	¿click?	publicidad online

Lo que puede hacer...

- En aprendizaje supervisado
 - lo que puedes hacer en **≤ 1 segundo...**
 - evaluar prefactibilidad con un **ante-proyecto**
- Es más sencillo cuando
 - concepto *simple* para aprender
 - muchos datos disponibles ($X \rightarrow Y$)

email	¿spam?
audio	transcripción a texto
Inglés	Español
imagen de un teléfono	¿defecto?
imagen, radar, gps	posición de otros autos
banner, dato usuario	¿click?



Optimización

Matemática aplicada que busca los valores que **minimizan una función** objetivo contemplando el precio computacional

Fuente: -

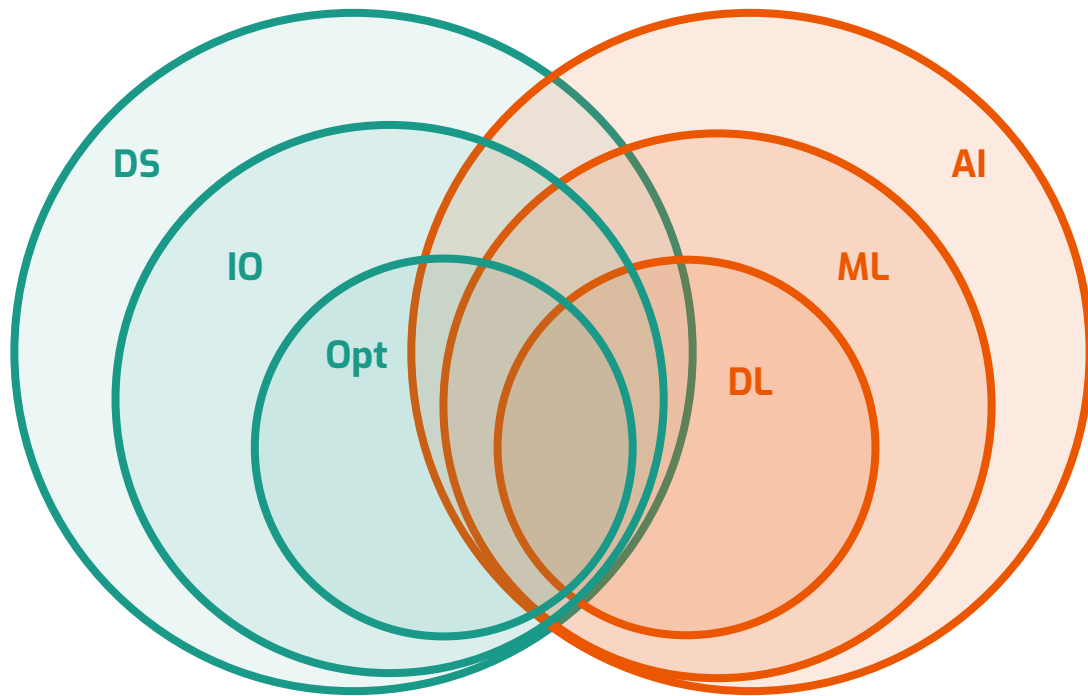




¿Por qué nombrar la optimización hoy?

- En **Data Science**
 - habitualmente es parte del análisis predictivo
 - siempre usada en análisis prescriptivo
- En **Artificial Intelligence**
 - en realidad aprender es optimizar
 - y las otras técnicas también

Intentemos poner estos conceptos juntos...



- dibujo controversial
- hay muchos más términos relevantes...



03

IMPLANTACIÓN EN LA EMPRESA

Impacto en Telcos y otras industrias

**Basado en ideas
y curso de:**



LANDING AI

de **Andrew Ng**

disponible en Coursera

<https://www.coursera.org/learn/ai-for-everyone>

<https://landing.ai/ai-transformations/>

Sept, 2019



Transformación digital, luego transformación DS-AI

Tranformación digital

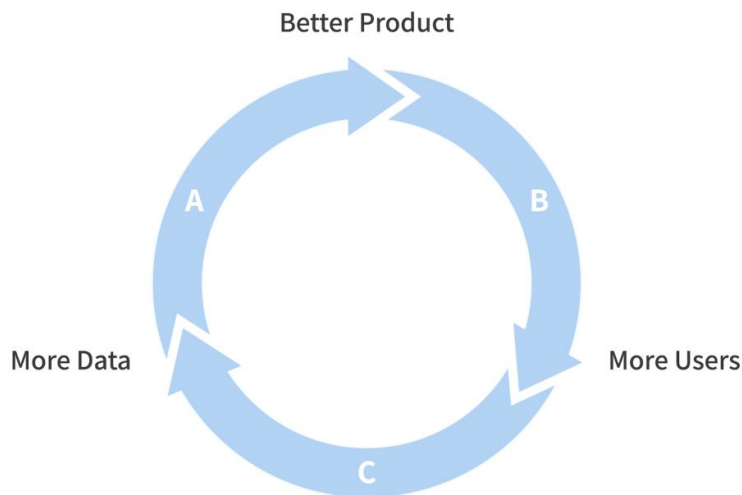
- La mejor tecnología sustenta los procesos, productos,...
- A/B testing
- Ciclo de vida de productos muy cortos
- Grandes decisiones tomadas por cargos técnicos especializados

Transformación DS-AI

- Adquisición estratégica de datos
- Repositorio único de datos
- Automatización generalizada
- Nuevos roles y distribución de tareas

Ciencia de datos aplicada a productos y servicios

The Virtuous Cycle of AI



Fuente: <https://landing.ai>

Crear productos más valiosos:

1. Usarla para crear una **ventaja competitiva** importante
2. Agregar muchas características, que hagan difícil a un competidor generar un producto igual
3. Alinear el esfuerzo con ciclo de retroalimentación positiva - **círculo virtuoso de la AI**
4. Alinear el esfuerzo con la **estrategia del negocio** (reducir costos, maximizar valor, aumentar efecto de red)



Pasos para la transformación DS-AI

- 1 Ejecutar proyectos pilotos para ganar impulso
- 2 Crear un equipo propio de DS-AI
- 3 Brindar amplia capacitación en DS-AI
- 4 Desarrollar una estrategia de DS-AI
- 5 Desarrollar comunicaciones internas y externas



Pasos para la transformación DS-AI

- 1 Ejecutar proyectos pilotos para ganar impulso**
- 2 Crear un equipo propio de DS-AI
- 3 Brindar amplia capacitación en DS-AI**
- 4 Desarrollar una estrategia de DS-AI
- 5 Desarrollar comunicaciones internas y externas



¿Cuándo empezar?

- Lo **antes posible**...
- Se puede avanzar con los **RRHH disponibles y tercerizar**
- Se puede avanzar con **pocos datos**. Ayuda en:
 - elegir mejores datos
 - elegir tecnología, modelos
 - cuantificar impacto
 - medir esfuerzo

(si se tienen muchos datos por lo general ayuda)



Capacidades necesarias

- tecnología y sistemas
- programación
- probabilidad y estadística
- lenguaje científico
- aprendizaje automático





Capacidades necesarias

- tecnología y sistemas
- programación
- probabilidad y estadística
- lenguaje científico
- aprendizaje automático



varios roles y herramientas necesarios para cubrir capacidades ...



Amplia capacitación en DS-AI

Gerentes y líderes de negocio

4 horas

- ¿cómo puede ayudar DS-AI a la empresa?
- ¿qué debe incluirse en la estrategia de desarrollo de DS-AI?
- ¿qué recursos se necesitan?

Jefes y gestores de proyectos

12 horas

- priorizar proyectos más valiosos y más factibles
- ¿cómo evaluar los resultados de un proyecto DS-AI?
- ¿cómo se hacer su seguimiento?
- ¿qué recursos se necesitan?

Ingenieros y personal técnico

≥100 horas

- desarrollar y desplegar software DS-AI
- ejecutar un proyecto DS-AI
- manipular datos

04

CASOS DE ESTUDIO

Ejemplos en Telcos y otras industrias

Diseño de red dorsal para transporte público

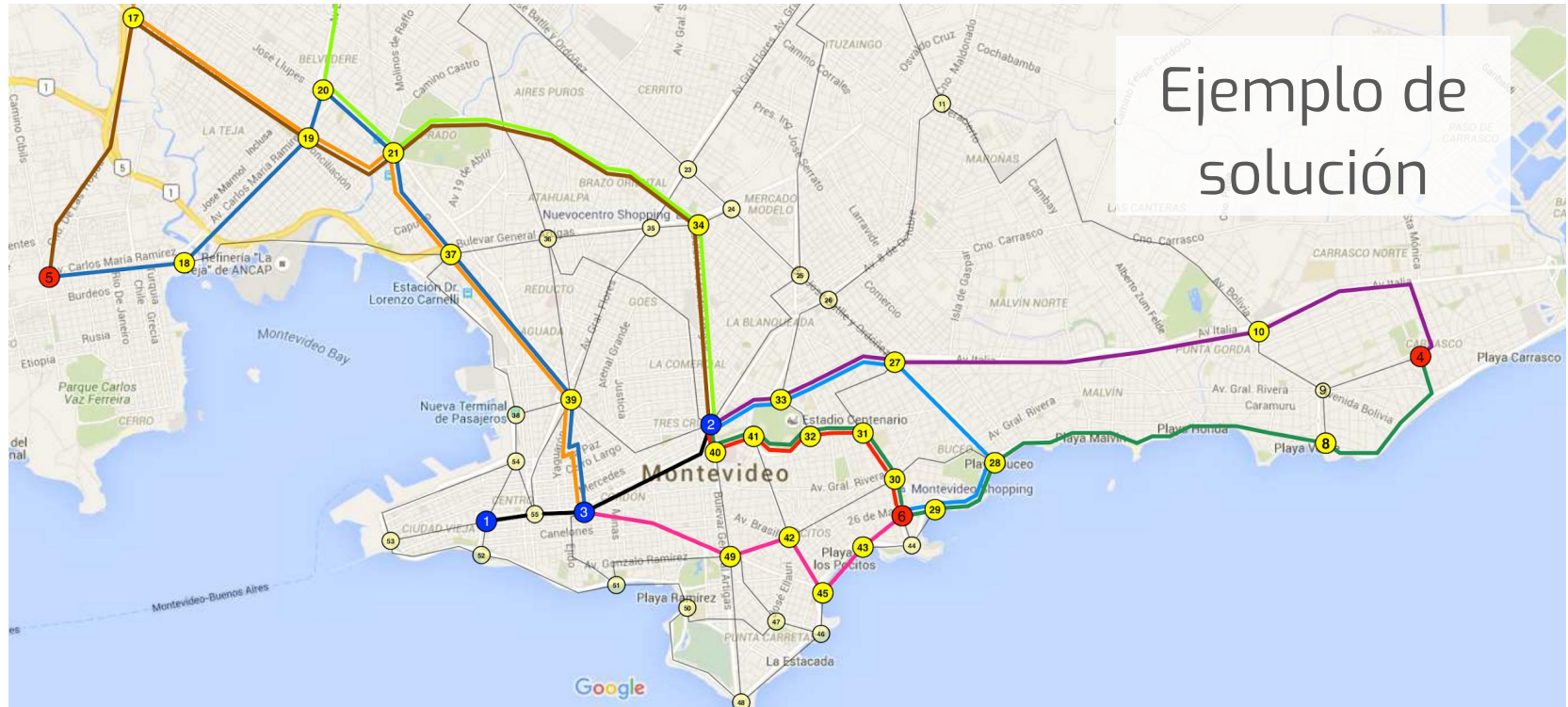
1/3

- definir una **red dorsal de tranvías** que reduce costo y nivel de servicio
- datos de boletos de STM
- **Costos** considerados
 - construcción
 - operación
 - usuario (tiempo)



Diseño de red dorsal para transporte público

2/3



Fuente: G. Faller , "Diseño Optimizado de un Backbone de Transporte Metropolitano para Montevideo". Actividad Integradora, Licenciatura en Computación, Instituto de Computación, Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay, 2019.

Diseño de red dorsal para transporte público

3/3

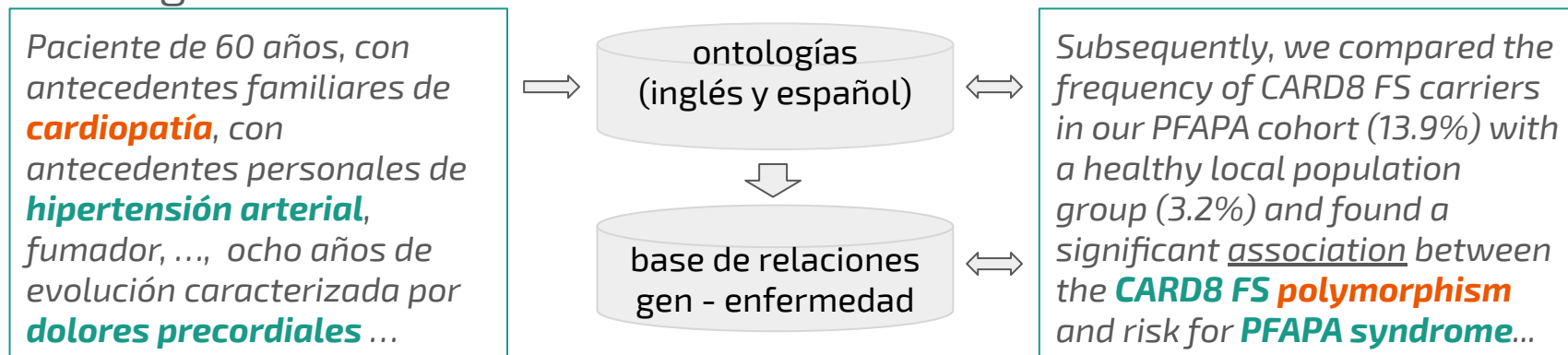
- problema de **optimización multiobjetivo**
- solución exacta y usando **algoritmos genéticos**
- resultado muy preliminar
 - notoria mejora en el nivel de servicio
 - económicamente viable
 - rediseño de los recorridos de los ómnibus

Origen	Destino	Ahora	Nuevo
Cerro	Carrasco	1h 35m	33m
Cerro	Centro	1h 10m	17m
Colón	Centro	1h 40m	19m
Colón	Pocitos	1h 20m	29m
Pocitos	Centro	25m	9m

Ejemplo de resultados preliminares.

Extracción de información de historias clínicas y artículos científicos para recomendación de estudios genéticos

- identificación de entidades en historias clínicas en español (uruguayas).
- extracción de entidades y relaciones a partir de artículos científicos
- cruzamiento de información utilizando bases de relaciones y ontologías bilingües



Cooperación de múltiples robots para la exploración

- cubrimiento de forma autónoma de una región desconocida por parte de un sistema robótico
- usos
 - exploración extraplanetaria o submarina
 - búsqueda y rescate, vigilancia,
 - patrullaje, agricultura, limpieza
- problemas subsidiarios
 - cartografía, localización, planificación de movimientos/tareas/roles/redes



Localización y construcción de mapas simultánea

- técnica utilizada por los robots para crear un mapa dentro de un entorno desconocido y al mismo tiempo hacer el seguimiento de su posición actual
- retroalimentación con la neurociencia experimental



Predicción de NDVI de cultivos agrícolas

1/2

- **NDVI** - Normalized Difference Vegetation Index
 - medición remota
- Relacionada con variables agronómicas de **relevancia**
 - índice de área foliar
 - déficit de nutrientes
 - afectación de pestes
 - **rendimiento del cultivo**

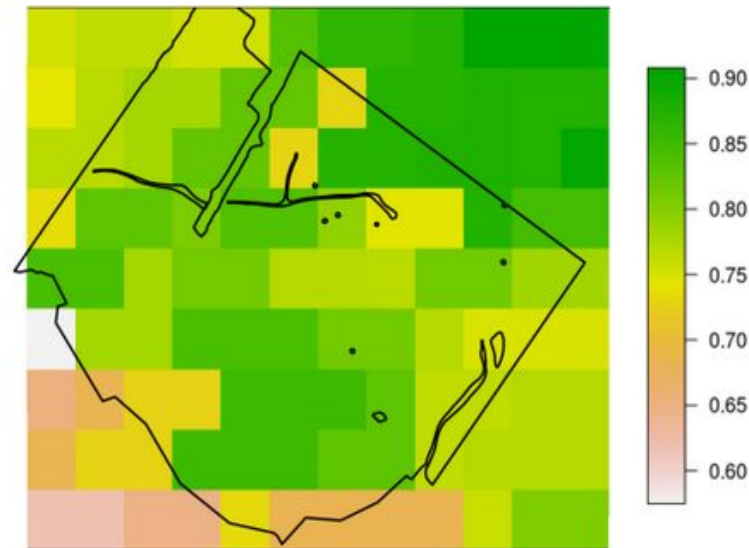
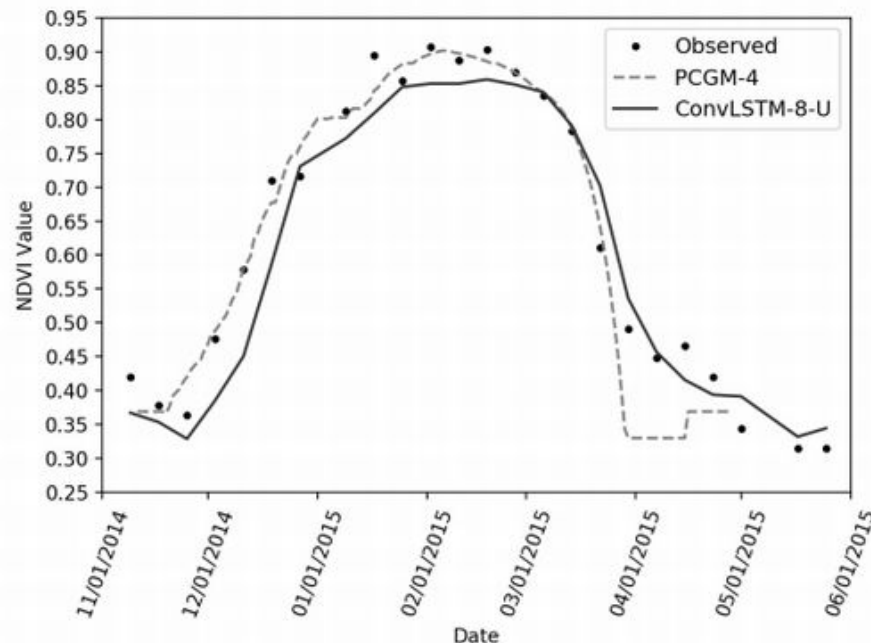


Imagen satelital de una chacra plantada con soja.
Tamaño de píxel: 250m² Fecha: 25/01/2015

Predicción de NDVI de cultivos agrícolas

2/2

- predecir futuras medidas de **NDVI** en base a información previa
- solución
 - sistema experto
 - red profunda convLSTM
 - error global (RMSE) del 7%



Evolución del NDVI en un píxel en la zafra 2014/2015

Anonimización de sentencias judiciales

- (semi)automatización del proceso de curación manual antes de la publicación de sentencias judiciales: anonimización, clasificación, etc.

- **Cancelación: Reconocimiento de Nombres (NER)**

Que ninguno de los funcionarios (...) observaron algún tipo de conducta fuera de lugar del **Sr. XX**. Que la denuncia fue realizada por la **Sra. XX** (...) Que lo único admitido por el **Sr. XX** es que compró un chip a la **Sra. XX**, pero no por eso se lo puede acusar de abuso y mucho menos de violación.

- **Sustitución: NER + CORREFERENCIAS**

Que ninguno de los funcionarios (...) observaron algún tipo de conducta fuera de lugar del **Sr. AA**. Que la denuncia fue realizada por la **Sra. BB** (...) Que lo único admitido por el **Sr. AA** es que compró un chip a la **Sra. CC**, pero no por eso se lo puede acusar de abuso y mucho menos de violación.

cancelación
sencillo,
pierde el relato

sustitución
mantiene el relato, más
complejo

Anonimización escalable de datos masivos y su aplicación al área de analíticas de aprendizaje

- crear **herramientas de anonimización** que extiendan la plataforma Hortonworks Data Platform (HDP)
 - código abierto de ejecución distribuida
 - siguiendo patrón de "privacidad desde el diseño" (privacy by design)
- caso de aplicación: datos de **Plan Ceibal**
 - PAM (plataforma adaptativa de matemática)
 - CREA2 (red social interactiva, docentes y estudiantes)
 - registros de la red de comunicaciones
 - Uruguay Estudia (datos de asistencia)
 - CRM Ceibal (inventario de dispositivos)



Impacto de la interacción en el aprendizaje

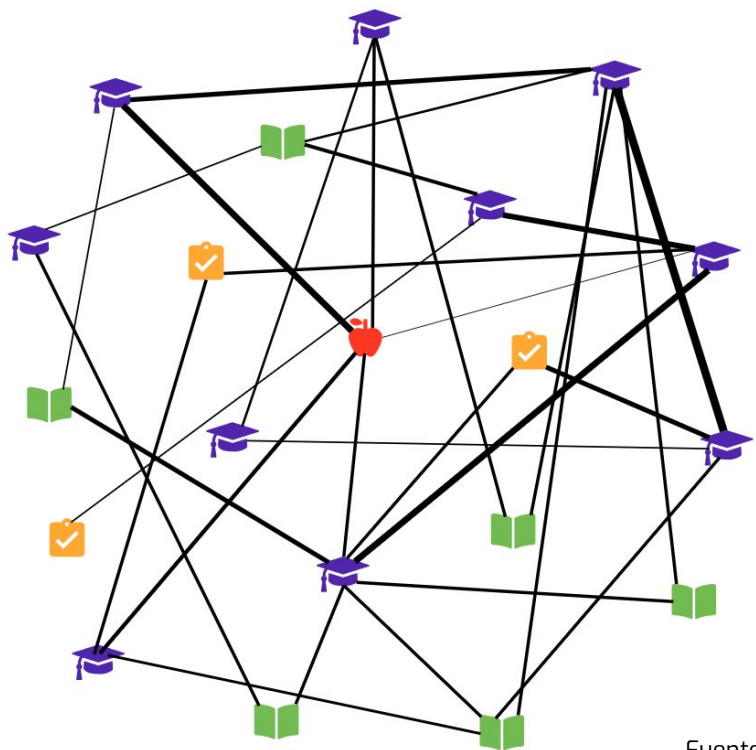
1/3

- integrar **perfil académico + social** del estudiante para descubrir patrones y producir **alertas/recomendaciones**
- datos de interacción en
 - aplicaciones de CEIBAL (CREA2, PAM)
 - Facebook
- aplicación de **análisis de redes + análisis de sentimientos**



Impacto de la interacción en el aprendizaje

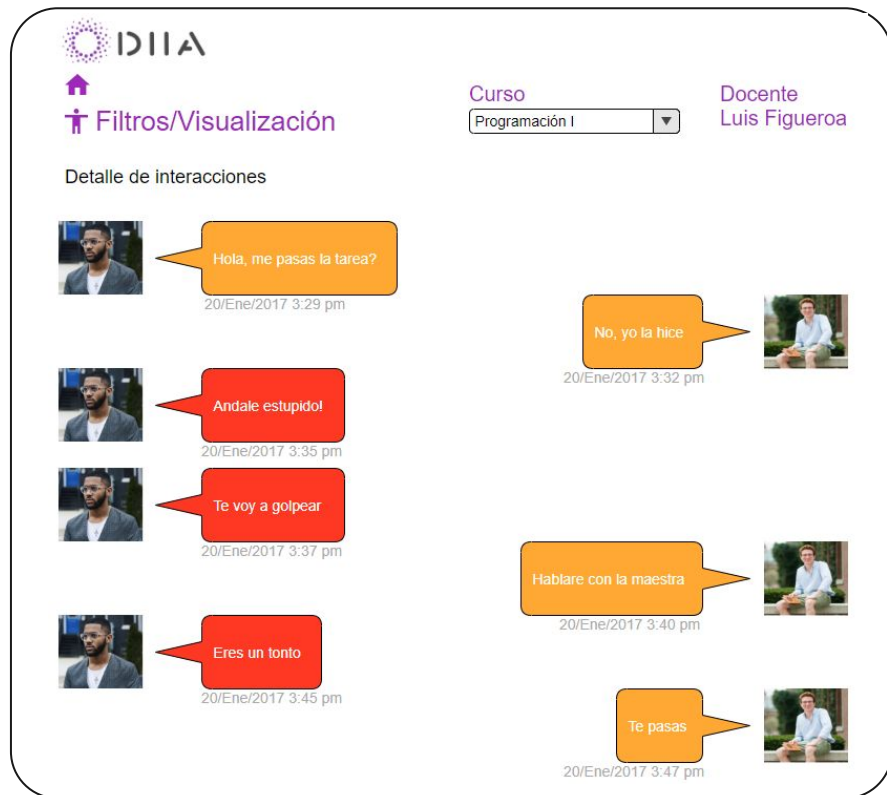
2/3



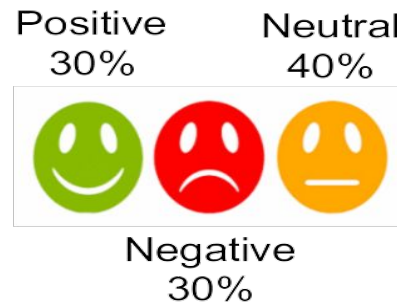
- métricas de análisis de redes
 - centralidad del estudiante
 - ...

Impacto de la interacción en el aprendizaje

3/3



- análisis de sentimientos
 - aplicado a mensajes y comentarios de estudiantes



Fuente: O. Cervantes et. al. "Use of Social Metrics to Discover Interaction Patterns that Impact Learning". Learning Analytics Latin America Workshop. Guayaquil, Ecuador, 2018.



Manos a la obra!

El próximo paso es el más importante

- Hablar con colegas sobre cómo avanzar
- Buscar capacitación
- Charlar con los varios expertos del tema disponibles
- Elegir y comenzar un proyecto sencillo
- Discutir sobre el impacto de avanzar en esta transformación...





gracias

Pablo Rodríguez Bocca



<https://www.linkedin.com/in/pablo-rodriguez-bocca/>

