



CONVOCATORIA PARA CONFORMAR EL BANCO DE RETOS ABIERTOS DE ENTIDADES PÚBLICAS Y EMPRESAS PRIVADAS, APLICANDO CIENCIA DE DATOS. 2020-2

1. PRESENTACIÓN

El Grupo Interno de Trabajo de la Estrategia de Talento de las Tecnologías de la Información - TI como parte de la Dirección de Desarrollo de la Industria TI del Ministerio de Tecnologías de la Información y las Comunicaciones, diseña e implementa programas que promueven el desarrollo del talento humano requerido en el sector TI y demás sectores económicos del país, contribuyendo a impulsar su competitividad y promoviendo el desarrollo de la economía digital en Colombia.

Tomando como base las acciones que adelanta la Estrategia de TI y teniendo en cuenta que las empresas a través de los años han recopilado una gran cantidad de información y datos sobre procesos, productos, servicios y clientes, dicha información al analizarla puede ser muy útil para la toma de decisiones, impactando en la prestación de mejores servicios o creación de mejores productos que satisfagan las necesidades de sus clientes.

En ese sentido, la Ciencia de Datos es utilizada junto con la inteligencia artificial combinada con diferentes habilidades estadísticas, informáticas y de conocimiento empresarial, para realizar un análisis de la información recopilada de diferentes fuentes, identificando variables, patrones de comportamiento y modelamientos de datos que contribuyen a que las empresas de diferentes sectores puedan tomar decisiones estratégicas, oportunas y efectivas. Estos análisis metodológicos son realizados por personas que, al tener las capacidades mencionadas, se les conoce como Científicos de Datos, áreas fundamentales para afrontar los retos de la cuarta revolución industrial.

Accenture Technology, líder mundial en mediciones de impactos de la tecnología, en su visión para el 2019, destaca a los trabajadores digitales como una de las tendencias tecnológicas afirmando que, a medida que la fuerza laboral se vuelve cada vez más humana con cada trabajador empoderado por sus habilidades y conocimiento, sumado a un nuevo y creciente conjunto de capacidades posibilitadas por la tecnología, por esta razón, es hora de que los negocios adapten las tecnologías que exitosamente crearon esta fuerza laboral preparada para el futuro y apoyen una nueva forma de trabajar en la era post digital, apoyándose en las tecnologías como big data y el procesamiento de datos.

Por otra parte, la empresa Thinknum, dedicada a la recopilación de información de Internet "Web scraping" para asesorar inversionistas manifiesta que, los Científicos de Datos se han convertido en una de las profesiones con más demanda y evidencia de ello, es que las 20 principales compañías más grandes de Estados Unidos, incluyendo empresas como Amazon, Booz Allen Hamilton, Walmart y Apple, están contratando Científicos de Datos.

Con el objeto de gestionar las capacidades de conocimiento tecnológico en inteligencia artificial y procesamiento de datos en beneficio de la productividad y competitividad del sector productivo de la industria, surge la necesidad de fortalecer y apoyar el desarrollo de competencias y habilidades digitales en el talento humano de la economía nacional, dirigidas a



estimular mayores oportunidades competitivas en las empresas y potenciar el ejercicio de comercialización y posicionamiento de productos y servicios basados en análisis de datos y de algoritmos predictivos, buscando con ello fortalecer a Colombia como líder en la región, lo cual también contribuirá a la generación de puestos de trabajo formales, con una de las mejores remuneraciones en los sectores de la economía.

Por lo anterior, MinTIC durante la vigencia 2020 se encuentra adelantando un proceso de formación a 1000 personas en Ciencia de Datos aplicada a la inteligencia artificial denominado: Data Science For All / Colombia ("DS4A"), a través de la cual se abordaron temáticas relacionadas con:

1. Herramientas de Big Data
2. Inferencia causal ML clásica
3. Modelos de agrupamiento Codificación de visualizaciones de datos Sesgos cognitivos 4. Recopilación de datos
5. Interpretación de datos
6. Transformación de datos
7. Visualización de datos
8. Reducción de dimensionalidad
9. Generación de hipótesis
10. Modelado lineal
11. Selección de modelo Optimización numérica
12. Bases de datos relacionales
13. Inferencia estadística
14. Diseño de encuestas y experimentos

Este proceso de formación lo está llevando a cabo Correlation ONE, junto con tres Instituciones de Educación Superior con Acreditación Institucional de Alta Calidad. Correlation ONE es líder mundial en el diseño e implementación de procesos de formación empresarial en Ciencia de Datos orientado a la fuerza laboral de alto nivel. Cuenta con más de 250,000 científicos de datos y más de 600 alianzas con los principales empleadores y universidades del mundo en Estados Unidos, China, Reino Unido, Canadá, Brasil, México, Argentina, Colombia e Irlanda. Los programas de formación desarrollados por Correlation ONE han sido incluidos en varios rankings y con estos han ayudado a fomentar el ecosistema de talento de IA en múltiples regiones de todo el mundo, han ganado reconocimiento mundial, incluyendo un Estudio de caso de Harvard Business School.

Correlation ONE ha diseñado programas sofisticados para empresas como: SoftBank, MinTIC, Citadel, JP Morgan, Point72, Two Sigma, MarketAxess, Twitch, ZS, Lyft, Deloitte, y la National Science Foundation y cuenta con un equipo de trabajo liderado por un profesor titular de Estadística en la Universidad de Harvard.



El modelo de formación utilizado por Correlation ONE es 40% teórico y 60% práctico. La parte práctica se desarrolla sobre casos reales o retos del sector público y privado, a través de los cuales se aplican los conocimientos y se brindan alternativas o soluciones a los requerimientos que surgen de sus procesos de transformación digital o cambios en su oferta de valor.

Durante el programa de formación, los participantes serán organizados en equipos, y cada equipo trabajará en el desarrollo de un reto específico. El reto es un desafío de ciencia de datos del mundo real de un sector particular. Con la dirección del equipo de enseñanza de nivel mundial, los participantes aplicarán sus aprendizajes durante el curso para desarrollar y dar solución a los retos seleccionados.

Por esta razón, MINTIC está interesado en conformar un banco de retos de entidades tanto públicas como privadas para ser desarrollados por los beneficiarios de este proceso de formación, y que los desarrollos respondan a las necesidades de algunas entidades tanto del orden nacional como regional que requieran soluciones a temas específicos de aplicación de Ciencia de Datos e Inteligencia Artificial.

2. OBJETIVO DE LA CONVOCATORIA

Seleccionar hasta 100 retos del sector público y privado que permitan la exploración, identificación, análisis y visualización de la información presentada en datos estructurados y no estructurados para la toma de decisiones estratégicas frente a los escenarios de negocio propios de las empresas y entidades públicas aplicantes, en el marco del programa de Ciencias de Datos.

3. CARACTERÍSTICAS DE LOS RETOS:

1. Deben ser problemas reales e identificados en cada entidad pública o empresa privada aplicante, que involucren el análisis de conjuntos de datos y que estén claramente definidos dentro del alcance de la ciencia de datos. Para conocer cómo identificar los problemas o retos que pueden ser resueltos aplicando ciencia de datos remítase a la infografía anexa "IDENTIFICACIÓN DE RETOS" y a los videos publicados en la convocatoria.
2. No se admitirán propuestas de desarrollo de Software. Los resultados serán un análisis exploratorio, identificación de variables, comportamientos y patrones que permitan el diseño y visualización de la información para toma de decisiones.
3. Los retos o casos deben contar con datos suficientes que sean relevantes para dar solución al problema planteado, para lo cual se firmará un acuerdo de confidencialidad (numeral 10), así mismo, deben estar disponibles para su consulta antes del 1 de agosto de 2020 y hasta el 15 de diciembre de 2020

4. REQUISITOS PARA PARTICIPAR

- Ser una organización pública o privada nacional o territorial, entidad descentralizada o adscrita a una entidad pública.
- Presentar los documentos de constitución (registro en cámara de comercio) en caso de ser empresas privadas, para entidades públicas bastará con mencionar en el anexo 1 la Ley, acuerdo o norma con el cual se crea. En todos los casos se debe adjuntar los



documentos de representación legal y documentos de identidad de quien ejerza esta condición para la entidad o empresa.

- Presentar la autorización del representante legal por escrito para postular el reto, ver Anexo 2.
- Designar por escrito a dos personas que pueden ser, funcionario y/o contratista y/o empleado de la entidad pública o empresa privada participante, uno para el rol de líder del reto y otro para el rol técnico, para que participen en las sesiones de trabajo programadas con el equipo técnico de la convocatoria.
- Sí la entidad pública o empresa privada presenta varios retos, cada uno debe contar con un líder de reto y un técnico designado.
- Disponer de la información necesaria cumpliendo con las características requeridas del equipo técnico en el marco de la convocatoria.
- El líder o técnico designado por la entidad o empresa debe brindar la información de manera oportuna a los equipos de estudiantes.

5. CRONOGRAMA

ACTIVIDAD DE LA CONVOCATORIA	FECHA INICIO	FECHA FIN
Apertura de la convocatoria	6 de julio de 2020	30 de julio de 2020
Revisión y selección de las propuestas	3 de agosto de 2020	21 de agosto de 2020
Publicación de resultados	26 de agosto de 2020	26 de agosto de 2020

6. PROCESO DE POSTULACIÓN

- Deben identificar los retos que consideren pueden ser resueltos a través de la aplicación de la ciencia de datos y que cuenten con la información necesaria para que los participantes de la convocatoria de ciencia de datos puedan desarrollarlos y brindarles una posible solución. Para llevar a cabo esta identificación de los retos deben revisar la infografía y los videos anexos, donde se dan pautas claras para lograr identificar que tipos de retos pueden ser desarrollados aplicando ciencia de datos.
- Enviar la cantidad de retos que deseen postular, para lo cual deben diligenciar un formulario para cada reto dentro de los plazos establecidos. La inscripción a partir del 6 de julio de 2020, debe realizarse a través del link: https://micrositios.mintic.gov.co/forms/2020/07_banco_elegibles_convocatoria_ciencia_datos/index.html
- Diligenciar los anexos técnicos 1 y 2 donde se debe describir detalladamente el problema a resolver y se debe brindar la autorización del representante legal de la entidad, para lo cual puede revisar los videos anexos.
- Los retos enviados después de la fecha y hora indicada en el calendario de la convocatoria y por medio distinto al establecido, no se tendrán en cuenta dentro del proceso de selección del banco de retos. Una vez inscrito no se podrán hacer modificaciones a los retos y formularios.



• 7. SELECCIÓN DE RETOS

Se seleccionarán los retos que estén inscritos y cumplan con todos los requerimientos definidos para formar parte del banco de retos abierto de entidades públicas y empresas privadas. Su selección o asignación en cuanto a desarrollo e implementación en el proceso de formación 2020, serán determinados por el equipo técnico (Correlation ONE) de la convocatoria, quien los evaluará y seleccionará teniendo en cuenta los siguientes elementos:

- Diligenciamiento del formulario en su totalidad.
- Envío de los documentos del representante legal.
- Disponibilidad de los datos e información relacionada con el reto.

Los retos seleccionados, serán notificados a las personas de contacto relacionadas en el formulario, con quienes se establecerá comunicación directa a fin de suscribir las alianzas necesarias y definir las políticas de derechos de autor que se apliquen a este caso particular.

Los retos que no cuenten con la información y datos necesarios para su desarrollo no serán seleccionados dentro del banco de retos.

8. DESARROLLO DE LOS RETOS

- Equipos de 5-7 participantes en el programa DS4A trabajarán en un reto durante 11 semanas. Aplicarán ciencia de datos y las técnicas de inteligencia artificial que aprenden durante el curso, con el asesoramiento de un equipo de instructores de clase mundial, para crear una solución impactante que pueda ser implementada por la entidad.
- Cada equipo contará con una persona de contacto, que será quien se comunicará con las entidades o empresas para solicitar información y llevar a cabo el desarrollo del reto.
- Cada equipo seleccionará un reto del banco de retos, por lo tanto, puede darse el caso de que más de un equipo elija un mismo reto. En este sentido, la entidad es quien decidirá con cuántos equipos tiene la capacidad de trabajar.
- Las entidades que sus retos sean seleccionados dentro del banco de retos, deberán firmar un acuerdo de confidencialidad con Correlation ONE y con cada equipo de DS4A que vaya a desarrollar su reto.
- Posteriormente, las entidades deberán remitir a los equipos que los seleccionaron toda la información y documentos necesarios para dar una solución al reto propuesto.
- Los equipos estarán en permanente contacto con las entidades y empresas a través de todo el proceso de desarrollo del reto.

9. LÍDER DEL PROYECTO

Es importante designar un punto de contacto en su empresa o entidad que pueda actuar como Líder o guía para los equipos del proyecto. Los equipos tendrán preguntas sobre los conjuntos de datos o sobre su negocio, y un mentor receptivo se asegurará de que aproveche al máximo sus equipos. También es una excelente manera de persuadir a los candidatos sobre los méritos de su organización, en caso de que desee contratar graduados de DS4A.

10. CONFIDENCIALIDAD



Todos los participantes de DS4A firmarán acuerdos de confidencialidad y no divulgación para proteger la privacidad de los datos. No se les permitirá usar los conjuntos de datos fuera del programa de capacitación. Si sus datos tienen información personal confidencial, es necesario anonimizar los campos para la protección de la privacidad y aplicación de la ley de protección de datos.

Así mismo, Correlation ONE firmará un acuerdo de confidencialidad con cada una de las entidades públicas o empresas privadas con el objetivo de proteger la información suministrada por las mismas.

11. RESULTADOS

- Cada equipo creará una aplicación totalmente funcional que resuelva el problema de ciencia de datos de la entidad. Como mínimo, la aplicación tendrá los siguientes tres componentes técnicos: interfaz de usuario front-end, canalización de datos, y herramientas & modelos analíticos. Los componentes específicos de la aplicación final dependerán en gran medida de los detalles específicos del proyecto.
- Informe de 10-20 páginas que detalla los métodos utilizados y los resultados del proyecto.
- Presentación de 10 minutos que resume el proyecto. El equipo presentará al resto del programa y excluirá cualquier información confidencial.
- Al finalizar la formación en ciencia de datos se llevará a cabo un evento para realizar networking con diferentes empresas del sector público y privado, con el objetivo de que puedan contratar a los graduados de DS4A.