

## Tarea 3: Enunciado y Proyectos

### 1. Administrativos

Tu grupo debe investigar sobre uno de los temas al reverso de la página, y presentar los resultados obtenidos en una presentación de 15 minutos, subida a mas tardar el 24 de Noviembre. Para los grupos armados (máximo 3 personas): Un responsable del grupo debe enviar un correo indicando el grupo y las primeras tres preferencias de proyectos (en orden) a Juan. Para las personas que no tienen grupo, pero que les gustaría uno: deben enviar un correo con las primeras tres preferencias de proyectos (en orden) a Juan. Se recomienda hacerlo a más tardar el viernes 06 de Noviembre a las 9:00, de lo contrario tu grupo perderá prioridad.

### 2. FAQ (o PF)

*¿Son los mismos grupos?*

No. Por defecto, espero que formen grupos donde no tengan a nadie en común con los grupos de la otra entrega. Sin embargo, puedo aceptar excepciones, aunque estas deben tener una cierta justificación.

*¿Todo igual que la vez pasada?*

Lo único que cambia es que ahora les voy a pedir que tengan sus videos el martes en la mañana, antes de las 10am. Esto para alcanzar a verlos todos el mismo día.

*Estamos atascados, nos cuesta decidir que hacer, o no sabemos como avanzar*

Para eso estamos. El próximo martes voy a estar conectado recibiendo consultas, la idea es que aprovechen para trabajar en esa hora. Pueden también fijar una reunión mandándole un correo a Juan. Opcionalmente, pueden organizar un discord si quieren ir compartiendo cosas entre ustedes y/o entre varios grupos.

### 3. Proyectos

**Nota:** Casi todos estos proyectos soportan varios grupos, pero ¡siempre y cuando no hagan exactamente lo mismo!

#### 3.1. Proyectos que involucran lectura de material literario, sintetización y presentación de un aspecto relevante

**IMPORTANTE:** Se debe buscar profundidad por sobre amplitud. Concentrarse en un punto específico siempre es mejor que dar una vista de pájaro del paper que sea tan por encima, que no le sirva a nadie.

**L1 In-database learning** . Hay dos buenas referencias en el syllabus de la página. Lo ideal es que se llegue a mostrar cómo se resuelve todo de forma que se optimiza el proceso de learning.

**L2. Fairness como un problema de datos.** Pueden mirar <https://arxiv.org/pdf/1908.07924.pdf> para un buen survey, y <https://arxiv.org/abs/1902.08283> es un buen ejemplo de una solución que se puede presentar (como reparar la base de datos para asegurar fairness).

**L3. Un tratamiento mas extenso de alguno de los dos framework para inferencia (potential outcomes o el via grafos)** Nuevamente, se sugiere concentrarse en un problema específico y mostrar como se logra tratar.

**L4. Ir más profundo en lo propuesto para CaRL.**

**L5. Descubrimiento de relaciones causales en bases de datos** . Pueden partir por [http://ceur-ws.org/Vol-2465/semex\\_paper4.pdf](http://ceur-ws.org/Vol-2465/semex_paper4.pdf)

#### 3.2. Proyecto sobre un ejercicio práctico de inferencia

**P1. Sobre el grafo de wikidata** Buscar el grafo de wikidata, especificar una pregunta causal (pensar p. ej. en el efecto de ser hombre en "largo de tu entrada de wikipedia", "numero de citas en tus papers", etc. )

**P2. Sobre el plebiscito** Buscar datos, armar de forma óptima el análisis visto en el video (explorar el resultado de distintos grafos causales)

#### 3.3. Proyecto sobre soluciones a BD en inferencias

**D1. Formalizar el proceso de encontrar las probabilidades y hacer inferencia de grafos como un problema de datos** . Similar a lo hecho para potential outcomes.

**D2. Extender lo realizado por ZaliQL para Neo4j** . Ojo con el problema visto de que la base de datos ya no es tan fácil de meter en una tupla (se podría parametrizar por alguna función de agregación).

**D3. Analizar lo hecho por ZaliQL o CaRL y mirarlo desde el punto de vista de Message Passing Networks, matrices o graph updates .**

### **3.4. Propone tu propia tarea**

Simplemente manden un mail a Juan para que quede confirmado el nuevo tema.