Using Random Fourier Features with Random Forests

Deliverable 3: Budget and Sustainability

Albert Ribes Marzá March 15, 2018

Contents

1	Self	Self-assessment on sustainability															3								
2	Analysis of the sustainability of the project															3									
	2.1	Enviro	on	mei	nta	ıld	lin	er	ısi	ioi	n														3
		2.1.1	F	PF	.																				3
		2.1.2	Ţ	Jse	ful	lif	e																		3
		2.1.3	F	Risk	s.																				4
	2.2	Econo	om	ic d	lim	en	sic	n																	4
		2.2.1	Ι	3ud	lge	t																			4
		2.2.2	A	Asse	ess	$m\epsilon$	$_{ m nt}$																		5
	2.3	Social	l di	$\mathrm{im}\epsilon$	ns	ior	ı																		5
		2.3.1	F	PF	.																				5
		2.3.2	Ţ	Jse	ful	lif	e																		5
		2.3.3	F	Risk	ζS.																				6

1 Self-assessment on sustainability

En 300 palabras, lo que he sacado de la encuesta

Poner todos los puntos fuertes y débiles

patata patata

2 Analysis of the sustainability of the project

2.1 Environmental dimension

2.1.1 PPP

El impacto sobre el medio ambiente a lo largo de la realización del TFG (consumo energético y generación de residuos).

El único consumo que hay es el uso del ordenador para todo.

No se genera ningún residuo.

2.1.2 Useful life

La huella ecológica que tendrá el proyecto durante toda su vida útil

Es un simple trabajo de investigación, no creo que tenga ninguna huella ecológica.

Si sale un algoritmo más eficiente, puede repercutir en el consumo que hagan las máquinas de otros para hacer machine learning.

2.1.3 Risks

El conjunto de eventualidades que podrían causar que el impacto ambiental del proyecto sea más negativo del previsto

Realmente no se me ocurre nada

2.2 Economic dimension

2.2.1 Budget

Hay que tener los campos:

• Costes directos por actividad

Indicar que todo el software que uso es libre, y por lo tanto no tengo costes de este estilo.

A pesar de que no tenga costes, quizá conviene hacer una tabla, y que para cada una ponga que tiene coste =0

El coste que ha tenido mi ordenador también entra aquí

Al final es coste humano y coste material, y el software está dentro de material

• Costes indirectos

El consumo de energía que ha tenido mi ordenador mientras hacía la investigación. Quizá también la luz que he gastado mientras hacía el proyecto.

El acceso a internet que he tenido que tener para hacer el proyecto.

Amortizaciones

Espero que el ordenador me dure 4 años

• Contingencias e imprevistos

Quizá se me estropea el ordenador

La estimación de costes se hace a nivel de actividades del Gantt, que son:

- Background approximation
- Decide the kernel
- Decide dimensionality
- Decide the changes
- Implement Fourier Mapping
- Get familiar with the module
- Modify the module
- Debug de code

- Find testing dataset
- Preprocessing
- Accuracy tests
- Time tests
- Study the results

Hay que estimar a cuanto serán esos costes y los mecanismos que tendré para evitar desviaciones

Para cada uno, recorrerme la lista de tareas que tengo hecha del deliverable anterior, y ver si tienen algo de esto. Poner la estimación de lo que creo que será

Direct costs

Las horas de trabajo, que se copian de la entrega anterior, y el consumo de energía del ordenador.

El consumo debe ser igual al de la PS4, que es 0'125 kWh cada hora

Indirect Costs

Depreciation

Unforeseen contingencies

2.2.2 Assessment

- Reflexionar sobre el coste que he estimado para la realización
- ¿Cómo se resuelven actualmente los problemas que quiere resolver mi proyecto?
- En qué mejora económicamente mi solución respecto de las otras existentes (en costes)

2.3 Social dimension

2.3.1 PPP

El impacto que ha tenido la realización del proyecto sobre las personas que han trabajado en él. Reflexionar sobre los cambios que la realización del proyecto ha provocado en mí y en mi entorno más directo.

A mí me está ayudando a aprender a hacer proyectos propiamente dichos, a planificarlos bien y todo eso. He tenido que organizarme mejor y todo eso.

2.3.2 Useful life

El impacto que tendrá la puesta en marcha del proyecto sobre los colectivos relacionados, ya sea de forma directa o indirecta.

Quizá puedo separar en dos secciones, en el caso que las conclusiones sean satisfactorias y en el caso de que no lo sean.

En el caso que este proyecto llegue a resultados satisfactorios, la comunidad científica dispondrá de una nueva técnica de machine learning con buenos resultados. También es posible que gracias a la realización de este proyecto, se puedan realizar otros productos que permitan hacer inferencia gastando menos recursos de energía, espacio de almacenamiento y tiempo de cálculo.

Si los resultados no son satisfactorios, es decir, las modificaciones que hacemos no parecen mejorar el rendimiento de los random forest, este es un estudio que lo comprueba, y por tanto evitará que otras personas vuelvan a hacer el mismo estudio, lo cual sería una pérdida de tiempo.

2.3.3 Risks

Eventualidades que que podrían causar que el impacto social que el proyecto sobre alguno de los colectivos relacionados con él sea más negativo del previsto

Realmente creo que no existe ningún riesgo con este proyecto. Tan solo aporta conocimiento al asunto.