**PMC**

De acuerdo a lo que hemos discutido hasta el momento el propósito de nuestro proyecto es inferir la probabilidad de éxito de iniciativas empresariales a partir del análisis de feeds sociales. Idealmente, nuestro proyecto debería tener una primera capa de análisis lingüístico cuyos resultados servirían de insumo a una segunda capa de análisis estadístico (en principio de clasificación).

El supuesto fundamental detrás de nuestro análisis es que existe una distribución de probabilidad conjunta subyacente sobre el espacio de predictores (incluyendo los “sociales”) y estados de la variable resultado que no se observa pero sobre la cual se puede hacer inferencia mediante la estimación de clasificadores derivados de algoritmos de aprendizaje.

Nos interesa entonces pensar en iniciativas empresariales (o cualquier tipo de fenómeno con relevancia comercial) que **a.** se lleven a cabo con regularidad y bajo condiciones estables en el tiempo **b.** exista a la fecha una cantidad importante de instancias etiquetadas sobre las cuales se puedan emplear algoritmos de entrenamiento y realizar pruebas de validación **c.** se vayan a observar nuevas instancias en el futuro cercano **d.** existan razones plausibles para pensar que las instancias observadas son iid…