

# Proposición de tema de tesis de Maestría

PUCP-GRPIAA Labs.

## Detección de metáforas en obras literarias

La metáfora es una figura retórica ampliamente utilizada en el lenguaje cotidiano (por ejemplo, “Tus cabellos de oro” o “los actores de la política de transporte”). Palabras que tienen una cierta acepción pueden ser comprendidas de una forma distinta a lo que expresa ese significado. Por ejemplo, algunos términos que tienen una connotación física, como golpear, pueden adquirir otro sentido cuando se lo usa metafóricamente, pudiendo dar a entender algo psíquico y no físico.

La metáfora es también ampliamente utilizada en las novelas literarias. La detección automática y la interpretación de metáforas es una tarea fundamental del Tratamiento Automático del Lenguaje Natural (TALN) y la recuperación de información.

El trabajo sobre la detección de metáforas han comenzado en la década de los 80s y existe una gran cantidad de técnicas sobre la estructura y los mecanismos del cómo estudiarlos. Los últimos años han sido testigo de un contraste salto tecnológico con un número cada vez mayor de los enfoques que explotan técnicas estadísticas. En comparación con los enfoques tradicionales en base a conocimientos codificados manualmente, existen métodos recientes que tienden a tener una cobertura más amplia; que a menudo son menos exactos, pero que son más eficientes y robustos. En este sentido, se han propuesto varios enfoques de detección de metáforas (ver referencias bibliográficas) y se basan en el uso de WordNet, treebank, etc. Existen estos recursos léxicos principalmente para Inglés, Francés y/o Español.

En este proyecto de tesis, nos interesamos el estudio de figuras retóricas, precisamente, en la utilización de métodos de aprendizaje automático para la detección e interpretación de metáforas en obras literarias (novelas) en lengua Española.

El alumno deberá proporcionar un estado del arte sobre los diferentes métodos de detección automática de metáforas, realizar un estudio comparativo de herramientas utilizables. El trabajo deberá finalizar con la construcción de un corpus de obras literarias en lengua Española, la prueba de los métodos y herramientas existentes en el corpus construido y a la proposición de algunas mejoras.

### Resumen:

- Título: Detección de metáforas en obras literarias.
- Tiempo de duración: 2 semestres académicos.
- Supervisor: Hugo Alatriza Salas.
- Lugar donde se realizará la tesis: Pontificia Universidad Católica del Perú.

### Referencias Bibliográficas

1. Matt Gedigian, John Bryant, Srinu Narayanan, and Branimir Cicic. Catching metaphors. Proceedings of the 3rd Workshop on Scalable Natural Language Understanding, pages 41–48. 2006.
2. Saisuresh Krishnakumaran and Xiaojin Zhu. Hunting elusive metaphors using lexical resources. Proceedings of the Workshop on Computational approaches to Figurative Language, pages 13–20. 2007
3. Shutova Lin Sun and Anna Korhonen. Metaphor identification using verb and noun clustering. Proceedings of the 23rd International Conference on Computational Linguistics, pages 1002–1010. 2010.

4. Yulia Tsvetkov, Elena Mukomel, Anatole Gershman. Cross-Lingual Metaphor Detection Using Common Semantic Features. First workshop on Metaphor in NLP (Meta4NLP). 2013.