

Precisión de la muestra en el aprendizaje de clases de conceptos definibles en la teoría de primer orden de los números reales.

José Luis Montaña

Universidad de Cantabria, España.

Dos de los objetivos principales de la Teoría del Aprendizaje tal y como se formaliza en los contextos más actuales, Teoría Estadística (STL) de V. Vapnik y Conocimiento Probablemente Aproximadamente Correcto (PAC learning) de Valiant son estudiar la complejidad (tamaño) de la muestra y discernir entre distintos modelos o clases de conceptos. En esta charla nos centramos en el primero de los problemas y estudiamos no sólo el número de ejemplos sino también la precisión (número de dígitos en la mantisa) con la que los mismos deben ser presentados en el caso de clases de aprendizaje formadas por conceptos definibles mediante fórmulas de primer orden sobre los números reales con grado y número de polinomios acotados.

La razón de plantear y estudiar el problema de la precisión surge de la reflexión de que en la práctica los fenómenos predictivos basados en ejemplos disponen de un conjunto de entrenamiento formado por muestras pseudoaleatoriamente generadas dentro de un rango finito de precisión mientras que los estudios teóricos centrados en el estudio del tamaño mínimo del conjunto de entrenamiento asumen precisión infinita en la generación aleatoria de muestras. Nuestras conclusiones intentan explicar por qué en la práctica es suficiente “poca precisión” (lineal en la precisión del error y en la precisión de la confianza) aunque puedan ser necesarios muchos ejemplos (exponencial en la precisión del error).