

**Título:** CONSENSUS POLICIES TO SOLVE BIOINFORMATIC PROBLEMS THROUGH BAYESIAN NETWORK CLASSIFIERS AND ESTIMATION OF DISTRIBUTION ALGORITHMS

**Nombre:** ARMAÑANZAS ARNEDILLO, RUBEN

**Universidad:** Universidad del País Vasco/Euskal Herriko Unibertsitatea

**Departamento:** Ciencia de la computación e inteligencia artificial

**Fecha de lectura:** 15/06/2009

**Mención a doctor europeo:** concedido

**Programa de doctorado:** Ingeniería Informática

**Dirección:**

> **Director:** PEDRO LARRAÑAGA MUGICA

> **Director:** IÑAKI INZA CANO

**Tribunal:**

> **presidente:** JOSÉ ANTONIO LOZANO ALONSO

> **secretario:** ANA ZUBIAGA ELORRIETA

> **vocal:** Antonio Bahamonde Rionda

> **vocal:** JOSÉ MARÍA PEÑA SANCHEZ

> **vocal:** YVAN SAEYS

**Descriptor:**

> INTELIGENCIA ARTIFICIAL

**El fichero de tesis** no ha sido incorporado al sistema.

**Resumen:** Los campos de la bioinformática y biología computacional son nuevas disciplinas de investigación nacidas del desarrollo que han sufrido las plataformas de análisis biológico. Hoy en día, estas plataformas generan una cantidad tan ingente de datos que tan solo mediante técnicas de aprendizaje automático y optimización dichos datos pueden ser explotados.

Dentro de estas áreas de conocimiento, se propone el uso de políticas consensuadas aplicadas a clasificadores Bayesianos y algoritmos de estimación de distribuciones. Mediante dichas nuevas propuestas se busca que los resultados obtenidos del análisis de datos bioinformáticos sean aún más robustos y fiables. Como aportaciones metodológicas, el trabajo incluye la presentación de una métrica de relevancia univariada consensuada, así como un método multivariado de selección de genes en experimentos de microarrays de ADN. Se ha desarrollado un algoritmo basado en remuestro estratificado con reemplazo, selección de variables y clasificadores Bayesianos para la creación de redes de interacción con peso de relevancia en sus arcos.

Como cuarta aportación, se presenta un consensado poblacional dentro de búsquedas mediante algoritmos de estimación de distribuciones que permite crear una jerarquía en la selección de variables mediante algoritmos de envoltura.

Todos los métodos propuestos en el trabajo han sido aplicados al análisis de diferentes tipos biológicos de datos, tales como microarrays de ADN, micro RNAs y espectrometría de masas. Además de verificar los métodos propuestos en bancos de pruebas amparadas por la comunidad científica, el análisis ha deparado resultados satisfactorios en patologías autoinmunes, cáncer colorrectal y esclerosis múltiple. El trabajo incluye así mismo la presentación de diferentes gráficas y representaciones visuales de resultados de gran utilidad para la propuesta de nuevas hipótesis de trabajo o validación en problemas del campo de la biología computacional.