



Gustavo Daniel Sosa Cabrera

Prof. M.Sc.

Nombre en citaciones bibliográficas: Gustavo Sosa-Cabrera

Sexo: Masculino

Nacido el 09-06-1980 en Asunción, Paraguay. De nacionalidad Paraguaya.

Datos del PRONII

Área: **Ingeniería y Tecnología - Activo**
 Categorización Actual: **Nivel Candidato a Investigador - Res.: 570/2022**
 Ingreso al PRONII: **Nivel Candidato a Investigador - Res.: 570/2022**

Información de Contacto

Mail: **gdsosa@pol.una.py**

Áreas de Actuación

- 1 Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Big Data - Datamining - Feature Selection
- 2 Ciencias Naturales, Matemáticas, Matemática Aplicada, Análisis Multivariante
- 3 Ingeniería y Tecnología, Ingeniería Eléctrica, Ingeniería Electrónica e Ingeniería de la Información, Ingeniería de Sistemas y Comunicaciones, Informática Médica - Usabilidad - Interoperabilidad Semántica

Formación Académica/Titulación

- 2018-En Marcha** Doctorado - Doctorado en Ciencias de la Computación
 Facultad Politécnica, Universidad Nacional de Asunción, Paraguay
 Tutor: Prof. Miguel García Torres, D.Sc. - Prof. Christian Schaerer Serra, D.Sc.
 Becario de: Consejo Nacional del Ciencia y Tecnología, Paraguay
 Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Macrodatos - Minería de Datos - Reducción de Dimensionalidad - Análisis Multivariante;
- 2008-2017** Maestría - Maestría en Ciencias de la Computación
 Facultad Politécnica Universidad Nacional de Asunción, Paraguay
 Título: Representatividad muestral en la incertidumbre simétrica multivariada para la selección de atributos, Año de Obtención: 2017
 Tutor: Prof. Miguel García Torres, D.Sc. - Prof. Santiago Gómez, M.Sc. - Prof. Christian E. Schaerer Serra, D.Sc.
 Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Minería de datos - Selección de Atributos - Medidas de correlación multivariante;
- 2007-2008** Especialización/Perfeccionamiento - Capacitación en Didáctica Universitaria
 Rectorado, Universidad Nacional de Asunción, Paraguay, Año de Obtención: 2008
 Áreas de Conocimiento: Ciencias Sociales, Ciencias de la Educación, Educación General, Educación Superior Universitaria;
- 1999-2005** Grado - Licenciatura en Análisis de Sistemas Informáticos
 Facultad Politécnica de la Universidad Nacional de Asunción, Paraguay, Año de Obtención: 2005
 Áreas de Conocimiento: Ingeniería y Tecnología, Ingeniería Eléctrica, Ingeniería Electrónica e Ingeniería de la Información, Ingeniería de Sistemas y Comunicaciones, Análisis, Diseño y Desarrollo de Software de Gestión;

Formación Complementaria

- 2010** Congresos Conferencia Latinoamericana de Informática (CLEI)
 Facultad Politécnica de la Universidad Nacional de Asunción, Paraguay
- 2021-2021** Cursos de corta duración
 Banco Interamericano de Desarrollo, Estados Unidos
 Título: IDB28x: El valor de la creatividad y la innovación: la Economía Naranja
 Horas totales: 30
- 2021-2021** Cursos de corta duración
 Banco Interamericano de Desarrollo, Estados Unidos
 Título: IDB6x: Gestión de proyectos de desarrollo
 Horas totales: 30

2020-2020	Cursos de corta duración Facultad Politécnica, Universidad Nacional de Asunción, Paraguay Título: Evaluación del aprendizaje en ambientes virtuales Horas totales: 10
2006	Encuentros Latinoware Parque Tecnológico Itaipú, Brasil
2008	Otros Programador Java ASEC Intermedio Fundación PROYDESA, Argentina
2012	Seminarios Escola de Verão do Laboratório Associado de Computação e Matemática Aplicada Instituto Nacional de Pesquisas Espaciais, Brasil
2016	Talleres Minería de Datos Facultad Politécnica de la Universidad Nacional de Asunción, Paraguay
2015	Talleres Introducción a los sistemas de información enfocados en el paciente Facultad Politécnica de la Universidad Nacional de Asunción, Paraguay
2015	Talleres Ingeniería del Software Basada en Evidencias Facultad Politécnica de la Universidad Nacional de Asunción, Paraguay
2014	Talleres Resolución de problemas de optimización combinatoria mediante técnicas de programación lineal entera Facultad Politécnica de la Universidad Nacional de Asunción, Paraguay
2014	Talleres Generación automática de test unitarios Facultad Politécnica de la Universidad Nacional de Asunción, Paraguay
2009	Talleres High Performance Computing Facultad Politécnica de la Universidad Nacional de Asunción, Paraguay
2009	Talleres Performance en la WWW Facultad Politécnica de la Universidad Nacional de Asunción, Paraguay

Idiomas

Inglés	Comprende: bien	Habla: bien	Lee: bien	Escribe: bien
Español	Comprende: muy bien	Habla: muy bien	Lee: muy bien	Escribe: muy bien
Guaraní	Comprende: bien	Habla: bien	Lee: bien	Escribe: bien
Portugués	Comprende: regular	Habla: regular	Lee: regular	Escribe: regular

Institución principal donde desarrolla sus actividades

Facultad Politécnica de la UNA - FPUNA

Actuación Profesional

Atenas Energía S.A. - Atenas Energía S.A.

Vínculos con la Institución

2014 - 2015 **Coordinador Técnico** C. Horaria: **5**

Actividades

10/2014 - 1/2015 Servicio Técnico Especializado
 Servicio realizado: Adecuación e implantación de un software libre de gestión documental en la estructura organizacional de una institución cliente

Centro de Investigación en Matemática - CIMA

Vínculos con la Institución

2017 - 2018 **Auxiliar de Investigación Junior** C. Horaria: **10**

Actividades

11/2017 - 10/2018 Proyecto de Investigación y Desarrollo
PINV15-706 Construcción de un modelo de incidencia de dengue aplicado a comunidades de Paraguay (COMIDENCO)
 Participación: Integrante del Equipo
 Descripción: Construir y verificar un modelo matemático para predecir la tasa de incidencia de dengue, en función a variables ambientales, sanitarias y epidemiológicas y en presencia de acciones de intervención anti-dengue
 Integrantes: Gustavo Sosa-Cabrera(Responsable)
 Situación: En Marcha; Tipo/Clase: Investigación.
 Alumnos:
 Financiadores: Consejo Nacional de Ciencia y Tecnología - CONACYT (Apoyo financiero)
 Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, ;

Centro Nacional de Computación de la UNA - CNC

Vínculos con la Institución

2010 - 2015 **Coordinador de Proyectos Informáticos** C. Horaria: **30**

Actividades

1/2010 - 3/2015 Servicio Técnico Especializado
Servicio realizado: Gestión de equipos de desarrollo en proyectos con tecnologías tanto anacrónicas como así también emergentes

2007 - 2009 **Líder de Equipo** C. Horaria: **30**

Actividades

1/2007 - 12/2009 Servicio Técnico Especializado
Servicio realizado: Análisis y diseño de sistemas. Coordinación de actividades de Implementación

2003 - 2006 **Desarrollador Back-End & Front-End** C. Horaria: **40**

Actividades

7/2003 - 12/2006 Servicio Técnico Especializado
Servicio realizado: Desarrollo de aplicaciones web y de escritorio con persistencia en bases de datos relacionales

Enving SRL - Enving SRL

Vínculos con la Institución

2012 - 2013 **Coordinador técnico** C. Horaria: **10**

Otras Informaciones: Proyecto de innovación tecnológica aplicado en meteorología

Actividades

7/2012 - 2/2013 Proyecto de Investigación y Desarrollo
PR-11-INN-41 Sistema de adquisición y gestión de datos meteorológicos
Participación: Coordinador o Responsable
Descripción: Desarrollo de un kit simple y auto-contenido de adquisición y transmisión de datos modular. Receptor de datos centralizados. Módulos de transmisión de datos. Software de manipulación, gestión, procesamiento y presentación de los datos. Formación de capacidades en el ámbito nacional para el ensamblado de instrumentos de medición en el área meteorológica.
Integrantes: Gustavo Sosa-Cabrera(Responsable)
Situación: Concluido; Tipo/Clase: Desarrollo.
Alumnos: Pregrado (6);
Financiadores: Consejo Nacional de Ciencia y Tecnología - CONACYT (Apoyo financiero)
Áreas de Conocimiento: Ingeniería y Tecnología, Otras Ingenierías y Tecnologías, Otras Ingenierías y Tecnologías, Informática - Electrónica - Meteorología;

Facultad Politécnica de la UNA - FPUNA

Vínculos con la Institución

2017 - 2018 **Personal Técnico** C. Horaria: **10**

Actividades

10/2017 - 3/2018 Proyecto de Investigación y Desarrollo
14-INV-471 Historial de salud personal único, confiable y ubicuo
Participación: Integrante del Equipo
Descripción: Proyecto cuyo objetivo es la de mejorar la organización y el acceso al historial de salud de pacientes a través de un sistema de historial de salud personal confiable, único y accesible ubicuamente
Integrantes: Gustavo Sosa-Cabrera(Responsable)
Situación: Concluido; Tipo/Clase: Investigación.
Alumnos:
Financiadores: Consejo Nacional de Ciencia y Tecnología - CONACYT (Apoyo financiero)
Áreas de Conocimiento: Ingeniería y Tecnología, Ingeniería Eléctrica, Ingeniería Electrónica e Ingeniería de la Información, Ingeniería de Sistemas y Comunicaciones, Informática en Salud;

2015 - Actual **Profesor Investigador** C. Horaria: **30**

Otras Informaciones: Núcleo de Investigación: Núcleo de Investigación y Desarrollo Tecnológico (NIDTEC).
Laboratorio Asociado: Laboratorio de Computación Científica y Aplicada (LCCA).
Grupo de Investigación: Computación Científica y Matemática Aplicada (CCyMA).
Línea de Investigación: Feature Selection based on Intercooperation.

Actividades

12/2015 - Actual Líneas de Investigación, Computación Científica y Matemática Aplicada, Núcleo de Investigación y Desarrollo Tecnológico
Selección de Atributos basados en Intercooperación
 Participación: Integrante del Equipo
 Descripción: Resumen de Actividades
 (*) Investigación teórica en la línea de Feature Selection aplicado a las tareas predictivas del Datamining.
 (*) Investigación aplicada en las líneas del Data Science y Big Data.
 (*) Desarrollo de Software de soporte y experimentación para el Data Science y Big Data.
 (*) Tutoría a alumnos de grado de la cátedra de Proyecto I y II de la carrera de Licenciatura en Ciencias Informáticas.
 Integrantes: Gustavo Sosa-Cabrera; Miguel García-Torres; Santiago Gómez; Christian Schaerer; Federico Divina;
 Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Análisis Multivariante;

2008 - Actual **Profesor Adjunto (Grado)** C. Horaria: 6

Otras Informaciones: Cátedras de Paradigmas de la Programación, Algorítmica II y Sistemas Operativos de la Carrera de Licenciatura en Ciencias de la Informática.

Actividades

7/2021 - Actual Docencia/Enseñanza, Licenciatura en Ciencias de la Informática
 Nivel: Grado
 Disciplinas dictadas:
 -Profesor Adjunto de Sistemas Operativos

2/2013 - Actual Docencia/Enseñanza, Licenciatura en Ciencias de la Informática
 Nivel: Grado
 Disciplinas dictadas:
 -Profesor Adjunto Paradigmas de Programación

8/2009 - Actual Docencia/Enseñanza, Licenciatura en Ciencias de la Informática
 Nivel: Grado
 Disciplinas dictadas:
 -Profesor Adjunto de Algorítmica II

2/2010 - 7/2011 Docencia/Enseñanza, Licenciatura en Ciencias de la Informática
 Nivel: Grado
 Disciplinas dictadas:
 -Encargado de la Cátedra de Algorítmica I

8/2010 - 1/2011 Docencia/Enseñanza, Licenciatura en Ciencias de la Informática
 Nivel: Grado
 Disciplinas dictadas:
 -Programación de Aplicaciones en Red

8/2009 - 1/2011 Docencia/Enseñanza, Ingeniería en Informática
 Nivel: Grado
 Disciplinas dictadas:
 -Auxiliar de la Cátedra de Lenguaje I

2/2008 - 7/2010 Docencia/Enseñanza, Licenciatura en Análisis de Sistemas Informáticos
 Nivel: Grado
 Disciplinas dictadas:
 -Auxiliar de la Cátedra de Programación I

2/2008 - 7/2010 Docencia/Enseñanza, Ingeniería en Electrónica
 Nivel: Grado
 Disciplinas dictadas:
 -Auxiliar de la Cátedra de Algoritmo

8/2009 - 1/2010 Docencia/Enseñanza, Programación de Computadoras
 Nivel: Grado
 Disciplinas dictadas:
 -Auxiliar de la Cátedra de Estructura de Datos

2/2009 - 7/2009 Docencia/Enseñanza, Programación de Computadoras
 Nivel: Grado
 Disciplinas dictadas:
 -Encargado de la Cátedra de Algoritmos II

Hospital de Policía Rigoberto Caballero - HPRC

Vínculos con la Institución

 2017 - 2017 **Consultor Informático**

 C. Horaria: **10**

Otras Informaciones: Consultoría independiente

Actividades

 2/2017 - 10/2017 Servicio Técnico Especializado
 Servicio realizado: Conforme la publicación científica del Dr. Lawrence Weed, se desarrolló la Historia Clínica Electrónica Orientada a Problemas con la infraestructura LAMP. Se montó la infraestructura de servidores virtualizados con KVM-QEMU para la integración continua.

Instituto de Investigacion de Ciencias de la Salud, UNA - IICS, UNA
Vínculos con la Institución

 2009 - 2010 **Asesor Informático**

 C. Horaria: **10**
Actividades

 7/2009 - 12/2009 Servicio Técnico Especializado
 Servicio realizado: Asesoría para el análisis, diseño y desarrollo de sistema integrado de gestión para los laboratorios de análisis clínicos y especializados

Universidad Columbia del Paraguay - COLUMBIA
Vínculos con la Institución

 2010 - 2010 **Profesor de Grado**

 C. Horaria: **5**
Actividades

 7/2010 - 12/2010 Docencia/Enseñanza, Ingeniería en Informática
 Nivel: Grado
 Disciplinas dictadas:
 -Profesor de la Cátedra de Algoritmia II

Significado de su trabajo en el contexto de los principales problemas planteados en su área:

El gran avance de la tecnología acaecido en los últimos años en lo referente a la recolección y almacenamiento de datos, ha desencadenado en la proliferación de conjuntos de datos complejos, heterogéneos y de gran envergadura, lo cual representa un reto cada vez mayor para la extracción de conocimiento; un conocimiento el cual se espera sea valioso en la toma de decisiones en el ámbito correspondiente.

En la actualidad, los datos conforman bases de datos que pueden definirse como complejas, desde el punto de vista de gran volumen, alta dimensionalidad, heterogeneidad de tipos, credibilidad difusa y temporalidad variada en su generación, por mencionar solo algunos de los criterios que caracterizan a las bases de datos de hoy en día.

De este modo, para superar los problemas derivados de la alta dimensionalidad de los datos, la selección de atributos (también conocida como selección de variables o selección de características) ha atraído una atención significativa durante las últimas décadas convirtiéndose en uno de los puntos clave tanto en el reconocimiento de patrones como en el aprendizaje automático.

Aunque se ha llevado a cabo una amplia investigación en la reducción de la dimensionalidad y ya se han propuesto en la literatura una gran variedad de métodos de selección de atributos, un aspecto importante pero generalmente descuidado en los mismos son las posibles interacciones o asociaciones de sinergias entre las variables bajo estudio para el modelado de los datos.

En este contexto, mi línea de investigación comprende el estudio de aquellas posibles interacciones/asociaciones categóricas multivariantes y sus posibles usos ya sea para el montaje de nuevos métodos de selección de atributos o el mejoramiento de aquellos pre-existentes atendiendo las particularidades en la construcción de modelos predictivos y/o descriptivos para el análisis de datos en los espacios denominados de alta dimensionalidad.

Los principales aportes, en cuanto a mis trabajos de investigación se refiere; se enmarcan en el establecimiento de una línea propia del proceso de selección de atributos que se caracteriza en lo denominado como inter-cooperación y que buscan a su vez generar contribuciones a las crecientes áreas del aprendizaje automático de máquinas, la minería de datos y los datos a gran escala. Como tópicos de estos estudios, las múltiples aplicaciones derivadas comprenden campos como ser: el reconocimiento de imágenes y señales, la minería de textos, la detección de intrusos, la bio-informática, los negocios y finanzas entre otros.

En adición, mi área de actuación comprende otras formas no-tradicionales de la representación del conocimiento y el aprendizaje automático como son la inferencia declarativa mediante la Programación Lógica Inductiva y la Programación Funcional Lógica Inductiva principalmente como temas de interés en la tutoría de trabajos finales de grado ya que se encuentran estrechamente ligados con la Cátedra de Paradigmas de Programación que vengo impartiendo desde el año 2013 en la carrera de Licenciatura en Informática de la FP-UNA como Profesor Adjunto.

Producción Técnica

Productos tecnológicos

1 Gustavo Sosa-Cabrera **Multivariate Symmetric Uncertainty and Other Measurements, 2017.**

Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Minería de datos - Selección de Atributos - Medidas de correlación multivariante;

Referencias adicionales: Estados Unidos/; Medio: Internet.

Finalidad: Estimators for multivariate symmetrical uncertainty, total correlation, information gain and symmetrical uncertainty of categorical variables. Disponibilidad: irrestricta.

Institución promotora/financiadora: Facultad Politécnica de la Universidad Nacional de Asunción.

Observaciones: Software package developed through the R programming language and published in the CRAN.

Where:

* R is 'GNU S', a freely available language and environment for statistical computing and graphics which provides a wide variety of statistical and graphical techniques: linear and nonlinear modelling, statistical tests, time series analysis, classification, clustering, etc.

* CRAN is a network of ftp and web servers around the world that store identical, up-to-date, versions of code and documentation for R.

Cursos de corta duración dictados

1 Gustavo Sosa-Cabrera **Curso de Introducción al GNU - LINUX, 2011. (Extensión extracurricular)**

Referencias adicionales: Paraguay/Español; Medio: Otros.

Institución promotora/financiadora: Facultad Politécnica de la Universidad Nacional de Asunción

Participación: Docente. Lugar: Carrera de Licenciatura en Ciencias de la Informática. Ciudad: San Lorenzo.

Observaciones: Primera Capacitación Voluntaria para el Grupo de Software Libre de la FP-UNA GSL-FPUNA con más de 60 inscriptos de las carreras de Programación, Electrónica y Análisis de Sistemas Informáticos, todos son alumnos del primero al cuarto año; esta capacitación tuvo como objetivo: La instalación y configuración del sistema Operativo Linux. Con esto buscamos que los integrantes del grupo manejen de forma teórica/práctica los conceptos básicos, para luego conformar grupos de investigación en esta línea de trabajo. Con esto queremos ver la posibilidad que la FP-UNA lidere el trabajo de migración que el Plan Director en TIC quiere hacer en las empresas y ministerios Públicos del estado.

2 Gustavo Sosa-Cabrera **Curso de GNU - LINUX, 2011. (Extensión extracurricular)**

Referencias adicionales: Paraguay/Español; Medio: Otros.

Institución promotora/financiadora: Facultad Politécnica de la Universidad Nacional de Asunción

Participación: Docente. Lugar: ETyC 2011. Ciudad: San Lorenzo.

Observaciones: Curso taller desarrollado en el marco de la exposición tecnológica, científica y cultural que tiene lugar en la Universidad Nacional de Asunción, organizada anualmente por la Facultad Politécnica (FP-UNA) en el mes de septiembre, en ocasión del aniversario de fundación de la Universidad Nacional de Asunción

Producción Bibliográfica

Artículos publicados en revistas científicas

Artículos completos publicados en revistas arbitradas

1 Santiago Gómez; Inocencio E. Ortiz-Samudio ; Gustavo Sosa-Cabrera; Miguel García-Torres; Christian Schaefer; **(RELEVANTE) Measuring Interactions in Categorical Datasets Using Multivariate Symmetrical Uncertainty, Entropy, v. 24 f: 1, 2022.**

Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Interacción Multivariante;

Medio: Internet. ISSN/ISBN: 1099-4300

Palabras Clave: interaction; intrinsic interaction; categorical data; patterned data; multivariable correlation; gain in multiple correlation; multivariate symmetrical uncertainty;

Observaciones: Journal Rank: JCR - Q2.

Impact Factor: 2.524 (2020).

2 Gustavo Sosa-Cabrera; Miguel García-Torres; Santiago Gómez; Christian Schaefer; Federico Divina; **Understanding a Version of Multivariate Symmetric Uncertainty to assist in Feature Selection, Journal of Computational Interdisciplinary Sciences, v. 10 f: 1, 2019.**

Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Minería de Datos;

Medio: Internet. ISSN/ISBN: 2177-8833

Palabras Clave: feature selection; symmetrical uncertainty; multivariate prediction of response; high dimensionality;

3 Gustavo Sosa-Cabrera; Miguel García-Torres; Santiago Gómez; Christian Schaefer; Federico Divina; **(RELEVANTE) A multivariate approach to the symmetrical uncertainty measure: Application to feature selection problem, INFORM SCIENCES, v. 494, p. 1-20, 2019.**

Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Minería de Datos, Selección de Atributos, Análisis Categórico Multivariante;

Medio: Internet. ISSN/ISBN: 0020-0255

Palabras Clave: multivariate symmetrical uncertainty; mutual information; entropy; multivariate sample size; total representativeness; feature selection;
Observaciones: Journal Rank: JCR - Q1.
Impact Factor: 5.5 (2019).

Trabajos en eventos

Trabajos completos en anales de eventos

- 1 Wildo Gómez; Gustavo Sosa-Cabrera; María Elena García; Versatilidad representativa y transición hacia la persistencia per se de datos estructuralmente complejos con Haskell. In: Conferencia Ibero Americana Computación Aplicada, 2022 Lisboa, Portugal CIACA & CIAWI 2022. 2022.**
Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Paradigma Declarativo - Programación Funcional - Tipos Estructurados Complejos ;
Medio: Internet.
Palabras Clave: versatilidad representativa del conocimiento; complejidad estructural de datos; paradigma declarativo; programación funcional; persistencia de datos; haskell;
Observaciones: Presentado el 9.NOV.2022 en la sesión de Computación Aplicada.
- 2 Félix Gómez; Pedro González; Derlis Salinas; Gustavo Sosa-Cabrera; María Elena García; Genericidad de funciones: el quid para la incorporación de dominios en un Sistema Funcional Inductivo basado en Haskell. In: XXV Congreso Argentino de Ciencias de la Computación (CACIC), 2019 Córdoba, Argentina. Libro de Actas del CACIC 2019. 2019.**
Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Minería de Datos Relacional;
Medio: Internet. ISSN/ISBN: 978--987-
Palabras Clave: funciones genéricas; logs; dominio específico; inferencia funcional inductiva; conocimiento de fondo; haskell;
Observaciones: Presentado en el XVI Workshop Bases de Datos y Minería de Datos.
- 3 Andrés Maciel Cardozo; Gustavo Sosa-Cabrera; María Elena García; Ensamblado ad hoc de clasificadores para la detección de cáncer de mama usando Scikit-learn. In: XXV Congreso Argentino de Ciencias de la Computación (CACIC), 2019 Córdoba, Argentina. Libro de Actas del CACIC 2019. 2019.**
Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Minería de Datos;
Medio: Internet. ISSN/ISBN: 978--987-
Palabras Clave: cáncer de mama; clasificación de tumores; ensamblado de clasificadores; sesgo; varianza; bagging; boosting; scikitlearn;
Observaciones: Presentado en el XVI Workshop Bases de Datos y Minería de Datos.
- 4 Helena Belén Rivelli; Gustavo Sosa-Cabrera; José Luis Vázquez Noguera; Cynthia Emilia Villalba Cardozo; Registro de Vacunación Electrónico Pediátrico Integrado de Fácil Uso. In: XII Congreso Internacional de Informática en Salud - INFOLAC 2018, 2018 La Habana, Cuba 2018.**
Áreas de Conocimiento: Ingeniería y Tecnología, Ingeniería Eléctrica, Ingeniería Electrónica e Ingeniería de la Información, Ingeniería de Sistemas y Comunicaciones, Informática médica - Usabilidad - Integración de la información;
Medio: Internet.
Palabras Clave: usabilidad; registro electrónico de vacunación; evaluación de usabilidad; fácil uso; integrado;
Observaciones: <http://www.convencionsalud2018.sld.cu/index.php/convencionsalud/2018>
- 5 Juan Manuel Báez Acuña; Clara Anuncia Paredes Cabañas; Gustavo Sosa-Cabrera; María Elena García; Descubriendo reglas de asociación en bases de datos del sector retail usando R. In: XXIV Congreso Argentino de Ciencias de la Computación, 2018 Tandil Bs. As. - Argentina 2018.**
Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Minería de Datos;
Medio: Internet. ISSN/ISBN: 978--950-
Palabras Clave: minería de datos; reglas de asociación; retail; bases de datos transaccionales; herramienta r; apriori; eclat;
Observaciones: Presentado en el XV Workshop Bases de Datos y Minería de Datos (WBDDM).
- 6 Gustavo Sosa-Cabrera; Miguel García-Torres; Santiago Gómez; Christian Schaerer; Federico Divina; Understanding a Version of Multivariate Symmetric Uncertainty to assist in Feature Selection. In: Conference of Computational Interdisciplinary Science (CCIS 2016), 2016 Sao Jose dos Campos - Brazil 2016.**
Medio: Internet.
Observaciones: In this paper, we analyze the behavior of the multivariate symmetric uncertainty (MSU) measure through the use of statistical simulation techniques under various mixes of informative and non-informative randomly generated features. Experiments show how the number of attributes, their cardinalities, and the sample size affect the MSU. We discovered a condition that preserves good quality in the MSU under different combinations of these three factors, providing a new useful criterion to help drive the process of dimension reduction.

Resúmenes simples en anales de eventos

- 1 Santiago Gómez; Miguel García-Torres; Gustavo Sosa-Cabrera; Emilio G. Sotto-Riveros ; Christian Schaerer; Classifying dengue cases using CatPCA in combination with the MSU correlation. In: Entropy 2021: The Scientific Tool of the 21st Century, 2021 Porto, Portugal Session Entropy in Multidisciplinary Applications. 2021.**
Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Datamining - Feature Selection - Multivariate Analysis;
Medio: Internet.

- Palabras Clave: multivariate correlation; multivariate sample size; entropic measures; principal components analysis; dengue fever;
- 2 Santiago Gómez; Gustavo Sosa-Cabrera; Miguel García-Torres; Inocencio E. Ortiz-Samudio ; Christian Schaerer; Multivariate Symmetrical Uncertainty as a measure for interaction in categorical patterned datasets. In: Entropy 2021: The Scientific Tool of the 21st Century, 2021 Porto, Portugal Session Information Theory, Probability and Statistics. 2021.**
Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Datamining - Feature Selection - Categorical Analysis;
Medio: Internet.
Palabras Clave: multivariate correlation; interaction; multivariate sample size; entropic measures; patterned datasets;
- 3 Santiago Gómez; Gustavo Sosa-Cabrera; Miguel García-Torres; Christian Schaerer; Is it correlation or interaction?. In: III Encuentro de Investigadores de la Sociedad Científica del Paraguay, 2018 Asunción 2018.**
Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Análisis Multivariante;
Medio: Internet.
Palabras Clave: correlation; interaction; models;
Observaciones: In this work we describe the MSU multivariate correlation, provide application examples, verify properties as an estimator, and prove its mathematical properties related to metric spaces. Formalization of interaction is also presented. Finally, the gained capability to measure multivariate correlation and interaction opens up new roads for the formulation and/or confirmation of mathematical models.
- 4 Gustavo Sosa-Cabrera; Miguel García-Torres; Santiago Gómez; Christian Schaerer; Federico Divina; Sample representativity comes as an aid in feature selection. In: Séminaire Math AmSud - Mathematics and Statistics of Big Data, 2016 Montevideo - Uruguay 2016.**
Áreas de Conocimiento: Ingeniería y Tecnología, Ingeniería Eléctrica, Ingeniería Electrónica e Ingeniería de la Información, Ingeniería de Sistemas y Comunicaciones, Minería de Datos;
Medio: Internet.
Palabras Clave: feature selection; symmetrical measures; representativeness; extreme sample;
Observaciones: This presentation shows the importance of properly calibrating sample size when seeking to separate "the wheat from the weed" in feature selection. In feature selection for a classification model, an important step is to identify those attributes that are relevant for prediction as well as those that are irrelevant or redundant. Thus, an objective measure of interaction and correlation among features is a necessity. Multivariate Symmetrical Uncertainty (MSU) is an entropy-based measure that has been recently proposed to identify interaction among a group of features. However, MSU's behavior has been less than predictable when presented with different numbers of features, high cardinalities and sample sizes. Hence MSU's behavior requires more extensive experimental evaluation. Through several experiments, we study the effect of number of features, cardinality of the features and sample size on MSU. We show that sample size has a role in moderating the tendency of MSU values to increase with higher cardinalities, and propose an empirical expression that relates the sample size and the cardinality of the features. This relationship inspires the concept of total representativity of a sample. We then employ a chi-squared goodness of fit test showing a way to determine the minimum sample size that assures total representativity.

Resúmenes expandidos en anales de eventos

- 1 Gustavo Sosa-Cabrera; Miguel García-Torres; Santiago Gómez; Christian Schaerer; Federico Divina; Understanding a multivariate semi-metric in the search strategies for attributes subset selection. In: Congresso Nacional de Matemática Aplicada e Computacional (CNMAC), 2018 Campinas SP - Brasil 2018.**
Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Minería de Datos;
Medio: Internet.
Observaciones: In classification tasks, a feature (i.e. independent variable) is considered relevant, irrelevant or redundant according to the information contained about the class (i.e. dependent variable). Feature selection (FS) consists of finding the minimal set of relevant features such that the classification error is optimized. A feature selection method has three components: evaluation criterion definition (e.g. feature relevance), evaluation criterion estimation, and search strategies for feature subset generation. In this work, we perform several experiments to study the effect of feature group conformation strategies such as Sequential Forward Search (SFS) & Sequential Backward Search (SBS) while using Multivariate Symmetrical Uncertainty (MSU) as a reliable measure of the association of the group with the class. For such purpose we have benchmarked the n-bit parity problem and the checkerboard pattern. Results demonstrate the evolution of the semi-metric MSU as a function of the number of selected attributes, cardinality and sample size.
- 2 Gustavo Sosa-Cabrera; Miguel García-Torres; Santiago Gómez; Christian Schaerer; Federico Divina; Effect of Sample Representativeness in Multivariate Symmetrical Uncertainty for Categorical Attributes. In: Business Analytics in Finance and Industry, 2018 Santiago - Chile 2017.**
Áreas de Conocimiento: Ingeniería y Tecnología, Otras Ingenierías y Tecnologías, Otras Ingenierías y Tecnologías, Minería de Datos;
Medio: Internet.
Observaciones: Multivariate Symmetrical Uncertainty (MSU), based on information theory concepts, is studied under four factors affecting its behavior. This measure of multivariate correlation and collective interaction among categorical attributes is now completely tested and ready for use. Total representativeness is introduced as a worthy condition in a sample of one categorical attribute, and extended to the multivariate case. A procedure is derived to calculate the minimum sample size guaranteeing a totally representative sample with desired probability level. The empirical expression separately found to approximate m using cardinalities of features, is in agreement with the procedure.

Libros y capítulos de libros publicados

Capítulos de libros publicados

- Juan Manuel Báez Acuña; Clara Anuncia Paredes Cabañas; Gustavo Sosa-Cabrera; María Elena García; Discovering Association Rules Using R. A Case Study on Retail's Database. In: (Org.). Communications in Computer and Information Science, Springer, Cham, 2019, v. 995, p. 199-210, ISSN/ISBN: 978-3-030-20787-8**

Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Minería de Datos;

Medio: Internet.

ISSN/ISBN: 978--3-03

Palabras Clave: data mining; association rules algorithm; retail; transactional databases; r tool;

Observaciones: Gustavo Sosa Cabrera es el autor por correspondencia del capítulo de libro (Journal Rank: JCR - Q3 in 2019).

Evaluaciones

Tutorías/Orientaciones/Supervisiones

Concluidas

Tesis/Monografías de grado

- Felix Gómez, Pedro González, Derlis Salinas, - Cotutor o Asesor - Genericidad de funciones: el quid para la incorporación de dominios en un Sistema Funcional Inductivo basado en Haskell, 2019**

Tesis/Monografía de grado (Licenciatura en Ciencias Informáticas) , FP-UNA - Facultad Politécnica de la Universidad Nacional de Asunción, Paraguay

País: Paraguay / Idioma: Español

Palabras Clave: funciones genéricas; logs; dominio específico; inferencia funcional inductiva; conocimiento de fondo; haskell;

Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Minería de Datos Relacional;

Observaciones: Gómez, F., González, P., Sosa Cabrera, G., & García Díaz, M. E. (2019). Genericidad de funciones: el quid para la incorporación de dominios en un Sistema Funcional Inductivo basado en Haskell. In XXV Congreso Argentino de Ciencias de la Computación (CACIC)(Universidad Nacional de Río Cuarto, Córdoba, 14 al 18 de octubre de 2019).

- Andrés Maciel Cardozo, - Cotutor o Asesor - Ensamblado ad hoc de clasificadores para la detección de cáncer de mama usando Scikit-learn, 2019**

Tesis/Monografía de grado (Licenciatura en Ciencias Informáticas) , FP-UNA - Facultad Politécnica de la Universidad Nacional de Asunción, Paraguay

País: Paraguay / Idioma: Español

Palabras Clave: cáncer de mama; clasificación de tumores; ensamblado de clasificadores; sesgo; varianza; bagging; boosting; scikitlearn;

Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Minería de Datos;

Observaciones: Maciel Cardozo, A., Sosa Cabrera, G., & García Díaz, M. E. (2019). Ensamblado ad hoc de clasificadores para la detección de cáncer de mama usando Scikit-learn. In XXV Congreso Argentino de Ciencias de la Computación (CACIC)(Universidad Nacional de Río Cuarto, Córdoba, 14 al 18 de octubre de 2019).

- Wildo Gómez, Leonardo Giménez, - Cotutor o Asesor - Versatilidad representativa y transición hacia la persistencia perse de datos estructuralmente complejos con Haskell, 2019**

Tesis/Monografía de grado (Licenciatura en Análisis de Sistemas Informáticos) , FP-UNA - Facultad Politécnica de la Universidad Nacional de Asunción, Paraguay

País: Paraguay / Idioma: Español

Palabras Clave: versatilidad representativa del conocimiento; complejidad estructural de datos; paradigma declarativo; programación funcional; persistencia de datos; haskell;

Áreas de Conocimiento: Ingeniería y Tecnología, Otras Ingenierías y Tecnologías, Otras Ingenierías y Tecnologías, Informática;

- Juan Manuel Báez Acuña, Clara Anuncia Paredes Cabañas, - Cotutor o Asesor - Descubriendo reglas de asociación en bases de datos del sector retail usando R, 2018**

Tesis/Monografía de grado (Licenciatura en Ciencias Informáticas) , FP-UNA - Facultad Politécnica de la Universidad Nacional de Asunción, Paraguay

País: Paraguay / Idioma: Español

Palabras Clave: minería de datos; reglas de asociación; retail; bases de datos transaccionales; herramienta r; apriori; eclat;

Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Minería de Datos;

Observaciones: Acuña, J. M. B., Cabañas, C. A. P., Sosa-Cabrera, G., & García-Díaz, M. E. (2018, October). Discovering Association Rules Using R. A Case Study on Retail's Database. In Argentine Congress of Computer Science (pp. 199-210). Springer, Cham.

Otras Referencias

Premiaciones

- 2022 HONOR AL MÉRITO (nacional), Rectorado UNA**

Padrino de Honor de los Licenciados en Ciencias Informáticas de la Facultad Politécnica de la Universidad Nacional de Asunción - Promoción 2021 denominada "Prof. Mst. Abel Bernal Castillo" con 66 egresados.

- 2 **2021 HONOR AL MÉRITO (nacional), Rectorado, Universidad Nacional de Asuncion**
Padrino de Honor de los Licenciados en Ciencias Informáticas de la Facultad Politécnica de la Universidad Nacional de Asunción - Promoción 2019 denominada "Renegociación transparente de Itaipú, una causa Nacional" con 62 egresados.
- 3 **2018 HONOR AL MÉRITO (nacional), Rectorado, Universidad Nacional de Asuncion**
Padrino de Honor de los Licenciados en Ciencias Informáticas de la Facultad Politécnica de la Universidad Nacional de Asunción - Promoción 2017 denominada "Homenaje al Centenario de la Reforma de Córdoba" con 65 egresados.
- 4 **2017 HONOR AL MÉRITO (nacional), Rectorado, Universidad Nacional de Asuncion**
Padrino de Honor de los Licenciados en Ciencias Informáticas de la Facultad Politécnica de la Universidad Nacional de Asunción - Promoción 2016 denominada "Renovando la UNA unidos en la diversidad" con 69 egresados.

Presentaciones en eventos

- 1 **Otra - Workshop de Tesis de Doctorado en Ciencias Computacionales e Ingeniería, 2020, Paraguay**
Nombre: Workshop de Tesis de Doctorado en Computación e Ingeniería. Tipo de Participación: Expositor oral - Información Adicional: Tema: "Detección y Explotación de la Interacción Multivariante para la Selección de Atributos aplicado en la tarea de Clasificación de la Minería de Datos".
Nombre de la institución promotora: Facultad Politécnica de la Universidad Nacional de Asunción
- 2 **Congreso - O CNMAC (Congresso Nacional de Matemática Aplicada e Computacional) é o evento mais importante da SBMAC (Sociedade Brasileira de Matemática Aplicada e Computacional) e tem por objetivo congrega professores, pesquisadores e profissionais, de universidades, centros de pesquisa e empresas, das mais diversas áreas da Matemática Aplicada e Computacional, para divulgar e discutir os resultados (conclusivos e em desenvolvimento) das suas pesquisas e tomar ciência da produção científica nas principais instituições nacionais., 2018, Brasil**
Nombre: Congresso Nacional de Matemática Aplicada e Computacional. Tipo de Participación: Poster - Información Adicional: Presentación del póster con el título "Understanding a multivariate semi-metric in the search strategies for attributes subset selection" en la Sesión Técnica denominada "ST3: Computação Científica" en fecha 20 de setiembre de 2018 en el IMECC/UNICAMP de Campinas - SP.
Nombre de la institución promotora: Sociedade Brasileira de Matemática Aplicada e Computacional
Áreas de Conocimiento: Ciencias Naturales, Ciencias de la Computación e Información, Ciencias de la Computación, Minería de Datos;
- 3 **Encuentro - Jornadas de Informática en Salud con el objetivo de compartir experiencias y dar a conocer los avances de la informática en salud en Paraguay , 2017, Paraguay**
Nombre: I Jornada de Informática en Salud. Tipo de Participación: Expositor oral
Nombre de la institución promotora: Facultad Politécnica de la Universidad Nacional de Asunción
- 4 **Congreso - The Conference of Computational Interdisciplinary Science (CCIS 2016) aims to be a meeting place of researchers and students working in areas of science that use scientific computing in their research projects. It is an initiative of the Pan-American Association on Computational Interdisciplinary Sciences (PACIS) and has a periodicity of two years. Although there are other forums that discuss related topics, such as Applied Mathematics, Bioinformatics, and Computational Physics, the CCIS 2016 seeks, in an innovative way, a broader dialog, which is inherently inter- and multidisciplinary, where researchers from different areas can share their experiences and find solutions to their computational problems. , 2016, Brasil**
Nombre: 4th Conference of Computational Interdisciplinary Science. Tipo de Participación: Expositor oral - Información Adicional: Presentación oral en inglés del artículo "Understanding a Version of Multivariate Symmetric Uncertainty to assist in Feature Selection" en la sesión correspondiente al "SS2b: Uncertainty quantification, inverse problems, and data assimilation" el día 8 de noviembre de 2016.
Nombre de la institución promotora: Instituto Nacional de Pesquisas Espaciais
- 5 **Taller - Jornada de capacitación en archivología, 2014, Paraguay**
Nombre: Taller de Gestión documental y Archivo: Mesa de Entradas y Organización del Archivo de Gestión. Tipo de Participación: Expositor oral
Nombre de la institución promotora: Facultad de Odontología Universidad Nacional de Asunción
- 6 **Otra - La ETyC es la mayor exposición tecnológica, científica y cultural que tiene lugar en la Universidad Nacional de Asunción, organizada anualmente por la Facultad Politécnica (FP-UNA) en el mes de setiembre, en ocasión del aniversario de fundación de la Universidad Nacional de Asunción, 2013, Paraguay**
Nombre: ETyC 2013. Tipo de Participación: Expositor oral - Información Adicional: Título de la disertación: "Introducción a las Pruebas Unitarias"
Nombre de la institución promotora: Centro Nacional de Computación de la UNA
- 7 **Taller - Jornadas de capacitación en gestión documental electrónica, 2012, Paraguay**
Nombre: Taller del Sistema de Gestión Documental - GDoc. Tipo de Participación: Expositor oral - Información Adicional: Las jornadas se realizaron los días 7 de setiembre y 2 de noviembre de 2012
Nombre de la institución promotora: Facultad Politécnica de la Universidad Nacional de Asunción
- 8 **Taller - Jornada de capacitación en archivología, 2012, Paraguay**
Nombre: Taller de Gestión documental y Archivo: Mesa de Entradas y Organización del Archivo de Gestión. Tipo de Participación: Expositor oral
Nombre de la institución promotora: Facultad de Arquitectura, Diseño y Arte - U.N.A.

Jurado/Integrante

Otros tipos

- 1 **Gustavo Sosa-Cabrera Participación en comités de Departamento de Informática, Área de Algoritmos con 22 candidatos y 13 puestos.. Otras Concurso Público para el Acceso, Permanencia y Ascenso al Escalafón Docente de la Facultad Politécnica de la UNA, 2019, Paraguay/Español**
Otra participación (),
Áreas de Conocimiento: Ingeniería y Tecnología, Otras Ingenierías y Tecnologías, Otras Ingenierías y Tecnologías, Algoritmos;
- 2 **Gustavo Sosa-Cabrera Participación en comités de Colegios técnicos en informática. Otras XII Encuentro de Bachilleres Técnicos en Informática, 2014, Paraguay/Español**
Otra participación (),
Obs: Participación en calidad de jurado del certamen.
- 3 **Gustavo Sosa-Cabrera Participación en comités de Colegios técnicos en informática. Otras XI Encuentro de Bachilleres Técnicos en Informática, 2013, Paraguay/Español**
Otra participación (),
Obs: Participación en calidad de jurado del certamen.

Información adicional:

Perfil en el Google Académico: <https://scholar.google.com/citations?user=Wrjw2XYAAAAJ>

Indicadores

Producción Técnica 3

Productos tecnológicos	1
Software	1
Cursos de corta duración dictados	2
Extensión extracurricular	2

Producción Bibliográfica 16

Artículos publicados en revistas científicas	3
Completo en revistas arbitradas	3
Completo en revistas NO arbitradas	0
Trabajos en eventos	12
Completo	6
Resumen	4
Resumen expandido	2
Libros y capítulos de libros publicados	1
Capítulo de libro publicado	1

Tutorías 4

Concluidas	4
Tesis/Monografía de grado	4

Otras Referencias 15

Otros datos Relevantes	4
Presentaciones en eventos	8

Jurado/Integrante

3