

# **Grigorios D. Tsoumakas**

Assistant Professor

Department of Informatics

Aristotle University of Thessaloniki

## **Extensive Curriculum Vitae**

- 1. Personal Information**
- 2. Studies**
- 3. Scholarships**
- 4. Professional Experience**
- 5. Teaching Experience**
- 6. Research Projects**
- 7. Scientific Activities**
- 8. Academic Activities**
- 9. Research Interests**
- 10. Publications**
- 11. Citations**
- 12. Other Information**

Thessaloniki, June 2017



## 1. Personal Information

Date of birth: 28/3/1977  
Place of birth: Thessaloniki  
Marital status: Married, 4-year-old twins  
Military obligations: Fulfilled. Pvt Research & Informatics, 8/8/2005 - 8/8/2006  
Home address: P.O. Box 1823, Trilofos 57500, Thessaloniki  
Work address: Department of Informatics, 54124, Thessaloniki  
Phone number: +30 2310 998887 (work), +30 6977 444082 (mobile)  
Email: [greg@csd.auth.gr](mailto:greg@csd.auth.gr)  
URL: <http://users.auth.gr/greg>

## 2. Studies

### University degrees

2001-2005 PhD, Dept. of Informatics, Aristotle University of Thessaloniki, Greece. Thesis: *Machine Learning for the Fusion of Multiple Distributed Intelligent Systems*.  
1999-2000 M.Sc. in Artificial Intelligence, Division of Informatics, University of Edinburgh, U.K. Grade A-, Excellent (distinction). Thesis: *A Knowledge Based System for Rock Fracture Classification in Deep Boreholes*.  
1994-1999 Honours, Dept. of Informatics, Aristotle University of Thessaloniki, Greece. Grade 8.1 out of 10, Very Good. Thesis: *MapExpert: Connecting Expert Systems with Geographical Information Systems for Land Evaluation*.

### Foreign languages

*English*: Very good reading, writing and communicating, Cambridge Proficiency, Grade B, 1995.

*German*: Elementary reading, writing and communicating, Grundstufe, Gesamtnote Befriedigends, 1993.

## 3. Scholarships

2002-2004 Three scholarships for corresponding paper presentations in European Conferences from a) the European Coordinating Committee for Artificial Intelligence (ECCAI), b) the Hellenic Artificial Intelligence Society (EETN) and c) the KDNET network of excellence.  
2001-2002 Two scholarships for attending corresponding schools, a) from the European Committee (project HPCF-2000-00401) for the school entitled "Intelligent Data Analysis" in Palermo, Italy and b) from Alexander S. Onassis public benefit foundation for the school entitled "The Data Avalanche: Reducing Information Overload", in Heraklion, Greece.  
1999-2000 Scholarship from the Student Awards Agency for Scotland covering the tuition fees of the MSc in Artificial Intelligence.

## 4. Professional Experience

- 03/2013 - Today Assistant Professor, Department of Informatics, Aristotle University of Thessaloniki.
- 08/2007 - 3/2013 Lecturer, Department of Informatics, Aristotle University of Thessaloniki
- 2002 - 2004 Consultant for SMEs of the areas of Thessaloniki and Chalkidiki in adopting e-business practices. Programme "Go-Digital" funded by the Greek Ministry of Development.
- 2001-2003 Lecturer of seminars on C, C++, Microsoft Word/Excel in the Association of Greek Physicists, dept. of Central and Western Macedonia (~ 400 hours).
- 2001-2002 Lecturer in IEK Kalamaria and IEK Thessaloniki 2, for a total of 377 hours, in the specialties Industrial Software Technician, Tele-Informatics Technician and Informatics Applications Specialist (3 semesters).
- 01/2000 - 09/2000 Database and web programmer (part-time).  
"Dynamic Blue Ltd", Edinburgh, UK.
- 09/1998 - 06/1999 Database and web programmer, technical support, network administration (part-time).  
Information Technology Center, Aristotle University of Thessaloniki, Greece.
- 03/1998 - 12/1998 Programmer of a commercial multimedia application (part-time).  
"Digital Arts" (in collaboration with MLS Multimedia), Thessaloniki, Greece.
- 10/1996 - 01/1998 Information technology manager (part-time).  
"STAR FM 97.1" radio station, Thessaloniki, Greece.

## 5. Teaching Experience

- 2015-today Teaching together with colleagues Petros Nicopolitids and Dimitris Vrakas the undergraduate course Mobile Networks & Application Development, Dept. of Informatics, AUTH, Greece. I'm covering the Android development part.
- 2009-today Teaching the postgraduate course Advanced Machine Learning & Knowledge Discovery, Dept. of Informatics, AUTH, Greece.
- 2008-today Teaching the undergraduate course Object Oriented Programming, Dept. of Informatics, AUTH, Greece. From 2008 till 2012, I was teaching this together with colleague Panagiotis Katsaros.
- 2007-today Teaching the undergraduate courses Decision Theory and Systems (together with colleague Ioannis Vlahavas) and Web Information Systems (together with colleague Athena Vakali), Dept. of Informatics, AUTH, Greece.
- 2007-2015 Teaching together with colleague Dimitris Vrakas the undergraduate course Intelligent Autonomous Systems, Dept. of Informatics, AUTH, Greece.
- 2011-2013 Teaching together with colleague Athena Vakali the postgraduate course Web Information Systems, International Hellenic University, Greece.
- 2007-2012 Teaching the undergraduate course Multiagent Systems, Dept. of Informatics, AUTH, Greece.

Since the academic year 2010-2011, I have supervised 14 postgraduate theses. Since the academic year 2007-2008, I have supervised 29 graduate theses.

## 6. Research Projects

- *Title:* Semantic indexing and question answering in big data. *Starting date:* 3/2017. *Ending date:* 2/2018. *Budget:* € 120.000. *Source of funding:* Atypon Systems Inc. *Role in the research team:* Researcher.
- *Title:* Semantic indexing and question answering in big data. *Starting date:* 11/2015. *Ending date:* 10/2016. *Budget:* € 60.000. *Source of funding:* Atypon Systems Inc. *Role in the research team:* Researcher.
- *Title:* Semantic indexing and question answering in big data. *Starting date:* 1/2015. *Ending date:* 07/2015. *Budget:* € 30.000. *Source of funding:* Atypon Systems Inc. *Role in the research team:* Researcher.
- *Title:* Daily gas consumption forecasting. *Starting date:* 8/2014. *Ending date:* 12/2014. *Budget:* € 10.000. *Source of funding:* Gas Distribution Company of Thessaly. *Role in the research team:* Researcher.
- *Title:* Customer data analysis system. *Starting date:* 6/2014. *Ending date:* 12/2014. *Budget:* € 28.000. *Source of funding:* Diamantis Masoutis S.A. *Role in the research team:* Researcher.
- *Title:* Information System for productivity and customer service improvement in the Public Transportation Companies. *Starting date:* 4/2014. *Ending date:* 6/2015. *Budget:* € 191.304 (€ 66.900). *Source of funding:* Greek General Secretariat for Research and Technology. *Role in the research team:* Researcher.
- *Title:* Intelligent System in Supply Chain Monitoring and Optimization. *Starting date:* 6/2012. *Ending date:* 6/2014. *Budget:* € 338.797 (€ 27.000). *Source of funding:* Greek General Secretariat for Research and Technology. *Role in the research team:* Researcher, applying machine learning algorithms on sales/supply data for sales prediction.
- *Title:* Multilabel Classification Methods for Domains with Large Number of Labels. *Starting date:* 1/1/2009. *Ending date:* 31/12/2009. *Budget:* € 3805. *Source of funding:* Greek State Scholarships Foundation. *Role in the research team:* Principal investigator, leading a bilateral collaboration project with TU Darmstadt, Germany. The project led to the following publication: *Tsoumakas, G., Loza Mencia, E., Katakis, I., Park, S.-H., Fürnkranz, J. (2009) "On the Combination of Two Decompositive Multi-Label Classification Methods", Proc. ECML/PKDD 2009 Workshop on Preference Learning, Bled, Slovenia, pp. 114-129.*
- *Title:* I-NET (Interoperability Enterprise NetWork): Consortium on Interoperability, Flexibility and Adaptability of B2B Transactions across Enterprises. *Starting date:* 1/6/2007. *Ending date:* 31/10/2008. *Budget:* € 100.750 (€ 25.340). *Source of funding:* Greek General Secretariat for Research and Technology. *Role in the research team:* Researcher, applying data mining techniques on past business transaction data with the goal of building a model that predicts the credibility of enterprises.
- *Title:* SWIM: Development of an Intelligent System for the management of Semantic Web Services. *Starting date:* 1/11/2005. *Ending date:* 31/10/2008. *Budget:* € 138.240 (€ 88.664). *Source of funding:* Greek General Secretariat for Research and Technology. *Role in the research team:* Researcher, developing data mining methods for automated classification of Web services according to their application domain. Part of my work in this project is described in the following publication: *Katakis, I., Meditskos, G., Tsoumakas, G., Bassiliades, N., Vlahavas, I. (2009) "On the Combination of Textual and Semantic Descriptions for Automated Semantic Web Service Classification", Proc. 5th IFIP Conference on Artificial Intelligence Applications & Innovations (AIAI 2009), Thessaloniki, Greece, 23-25 April 2009.*
- *Title:* SWEL: Planning Systems in the Semantic Web for E-Learning. *Starting date:* 1/1/2005. *Ending date:* 31/12/2007. *Budget:* € 50.000. *Source of funding:* Greek Ministry of National Education and Religious Affairs. *Role in the research team:* Researcher, responsible for the design of a text classification system that matches the textual description of an educational resource to one or more nodes of an educational terms ontology. Part of my work in this project is described in the following publication: *Vrakas, D., Tsoumakas, G., Kokkoras, F., Bassiliades, N. Vlahavas, I., Anagnostopoulos, D. (2007) "PASER: A Curricula Synthesis System based on Automated Problem Solving", International Journal on Teaching and Case Studies (IJTCS), 1(1/2), pp. 159-170, Inderscience.*
- *Title:* EUROCITIZEN: An interactive, agent-based, on line edutainment environment to promote European intercultural education and cooperation among citizens of Europe. *Starting date:* 1/9/2002. *Ending date:* 28/2/2005. *Budget:* € 750.000 (€ 69.633). *Source of funding:* European Union under the

IST program. *Role in the research team:* Developer, responsible for developing the question selection mechanism.

## 7. Scientific Activities

### Invited talks and lectures

- Invited talk on “Exploiting Target Dependencies in Multi-Target Prediction” at the BENELEARN 2016 conference, Kortrijk, Belgium, September 12-13, 2016.
- Invited talk on “Multi-Target Prediction and Applications in the Publishing, Energy and Retail Industries” at the 2016 South-East Forum on Data Science, Belgrade, Serbia, June 20-21, 2016.
- Invited lecture on “Semantic Indexing of Biomedical Literature” at XEROX Research Center Europe, Grenoble, France, February 20, 2014
- Invited lecture on “Multi-Label Learning for Concept-Based Retrieval: The ImageCLEF 2011 Experience” at the Informatics and Telematics Institute, Centre for Research & Technology – Hellas, Thessaloniki, Greece, November 16, 2011.
- Invited postgraduate lecture on “Learning from Multi-Label Data” at the Department of Informatics of the Ionian University, Greece, Kerkyra, Greece, May 10, 2011.
- PASCAL2 invited talk on “Ensemble Methods for Multi-Label Data” at the 2010 Supervised and Unsupervised Ensemble Methods workshop, part of the ECML PKDD 2010 conference, Barcelona, Spain, September 20, 2010.
- Invited lecture on “Learning from Multi-Label Data” at the 1<sup>st</sup> Hellenic Artificial Intelligence Summer School (HAISS), Thessaloniki, Greece, August 24, 2009

### Prizes and Awards

- 1st place at the large-scale online biomedical semantic indexing challenge organized in 2013 by the European FP7 research project BioASQ, outweighing the state-of-the-art operational system of the National Library of Medicine and the approaches of several teams, including Mayo Clinic and the National Center for Biotechnology Information. In collaboration with Prof. Ioannis Vlahavas and US company Atypon. In 2014 and 2015 we achieved the 2<sup>nd</sup> place.
- 8th place at the bird classification challenge organized by the 9th IEEE international workshop on machine learning for signal processing (MLSP 2013) and hosted by Kaggle.
- 1<sup>st</sup> place at the concept-based retrieval subtask of the photo annotation task of ImageCLEF 2011. In collaboration with MSc student Kostantinos Sechidis and PhD student Eleftherios Spyromitros-Xioufis.
- 2<sup>nd</sup> place at the music instruments recognition track of the ISMIS 2011 international contest on music information retrieval.

### Conference Organization

- SETN 2015 workshops/tutorials chair.
- WISE 2014 challenge co-chair together with Apostolos Papadopoulos and Weining Qian.
- ECML PKDD 2012 publicity chair.
- ECML PKDD 2011 publicity chair together with Annalisa Appice.

### Workshop Organization

- "Big Multi-Target Prediction (BigTargets)" at the ECML PKDD 2015 conference in Porto, Portugal, together with Willem Waegeman, Krzysztof Dembczynski, Tapio Pahikkala, Antti Airola, Giorgio Valentini and Massih-Reza Amini.
- "Learning from Multi-Label Data (MLD'10)" at the ICML 2010 conference in Haifa, Israel, together with Min-Ling Zhang and Zhi-Hua Zhou.
- "Learning from Multi-Label Data (MLD'09)" at the ECML PKDD 2009 conference in Bled, Slovenia, together with Min-Ling Zhang and Zhi-Hua Zhou.
- "Mining Social Data (MSODA '08)" at the ECAI 2008 conference in Patras, Greece, together with Alexandros Nanopoulos.

### Tutorials

- "Mining Complex Dynamic Data" at the ECML PKDD 2011 conference in Athens, Greece, together with Myra Spiliopoulou, Irene Ntoutsi, Arthur Zimek and Hans-Peter Kriegel.

- "Learning from Multi-Label Data" at the ECML PKDD 2009 conference in Bled, Slovenia, together with Min-Ling Zhang and Zhi-Hua Zhou.

### **Editorial Board Member**

- Data Mining and Knowledge Discovery, Springer, 2016-today.

### **Special Issues**

- "Learning from Multi-Label Data", Machine Learning 88(1-2), July 2012, together with Min-Ling Zhang and Zhi-Hua Zhou.
- "ECAI'08 Workshop on Mining Social Data", Int. Journal on Data Warehousing and Mining 6(1), 2010, together with Alexandros Nanopoulos

### **Selected Journal Paper Review Invitations**

- *Journal of Machine Learning Research, Machine Learning Journal, Neurocomputing, Data Mining and Knowledge Discovery, IEEE Transactions on Neural Networks and Learning Systems, IEEE Transactions on Knowledge and Data Engineering, IEEE Transactions on Pattern Analysis and Machine Intelligence.*

### **Selected Programme Committee Memberships**

- Knowledge Discovery in Databases (KDD) 2016, 2015, 2011, 2010
- IEEE International Conference on Data Mining (ICDM) 2016, 2015, 2014
- SIAM International Conference on Data Mining (SDM) 2017, 2016, 2015, 2013
- European Conference on Machine Learning & Principles and Practice of Knowledge Discovery in Databases (ECML PKDD) 2016, 2015, 2014, 2013, 2012, 2011
- International Conference on Machine Learning (ICML) 2014, 2013, 2012, 2010
- Asian Conference on Machine Learning (ACML) 2017, 2014, 2012
- Discovery Science (DS) 2014, 2013
- International Joint Conference on Artificial Intelligence (IJCAI) 2017, 2015, 2013, 2011
- European Conference on Artificial Intelligence (ECAI) 2014, 2012
- AAAI Conference on Artificial Intelligence (AAAI) 2013, 2012
- Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD) 2009

## **8. Academic Activities**

### **Supervision of PhD Theses**

1. Yannis Papanikolaou, Text Mining, 03/2014 – today.
2. Eirini Papagianopoulou, Knowledge Discovery from Linked Open Data, 05/2014 – today.
3. Dimitris Dimitriadis, Text Mining and Natural Language Processing, 05/2015 – today.
4. Athanasios Lagopoulos, Web Usage Mining, 03/2017 – today.

### **External examiner of international PhD theses**

1. Daniel Hernandez-Lobato, "Prediction Based on Averages over Automatically Induced Learners: Ensemble Methods and Bayesian Techniques", Escuela Politécnica Superior, Universidad Autónoma de Madrid, Spain, January 2010.
2. Jesse Read, "Scalable Multi-Label Classification", Department of Computer Science, University of Waikato, New Zealand, September 2010.
3. Arnaud Joly, "Exploiting Random Projections and Sparsity with Random Forests and Gradient Boosting Methods", University of Liege, Belgium, March 2017.

### **Host of PhD students**

- Hosted PhD Student Newton Spolaor from the University of São Paulo at São Carlos for 5 months (4/2013 - 9/2013) as part of an accepted research proposal by the São Paulo Research Foundation – FAPESP.
- Hosted PhD Student Everton Alvares Cherman from the University of São Paulo at São Carlos for 5 months (6/2012 - 10/2012) as part of an accepted research proposal by the São Paulo Research Foundation – FAPESP.

- Hosted PhD Student William Groves from the University of Minnesota for 2 months as part of an accepted application to the IJCAI'11 call for European AI labs to host a US graduate student for an extended visit.

## 9. Research Interests

My research interests belong to the areas of Machine Learning and Knowledge Discovery in Databases. In specific, I have worked in the past/I am still working on the following topics:

- *Learning from Multi-Target Data*. Development of efficient and effective algorithms for predictive modelling from multi-target data (e.g. text, web, biology, music, image/video). Papers: J6, J14, J15, A6, C15, C19, C24, C26, C30, C31, C32, C34, C35, C37, C39, C47, C51, C52.
- *Ensemble Methods*. Development of algorithms for the combination of multiple predictive models with the aim of increasing predictive accuracy. Papers: J3, J10, J11, J13, A5, C2, C9, C14, C18, C20, C22, C33, C48.
- *Text Mining*. Development of algorithms for learning from textual data, with emphasis on text streams and their challenges (e.g. concept drift). Papers: J8, J12, A4, C11, C16, C17, C21, C43, C49.
- *Data Mining and Semantic Web*. Using semantic web technology to enhance data mining techniques, as well as using data mining techniques to learn from and improve the quality of the semantic web. Papers: C53
- *Learning for Planning*. Development of adaptive planning systems with the aid of machine learning techniques. Papers: J2, A3, C5, C6, C7, C8, C10.
- *Distributed Data Mining*. Development of algorithms that learn from geographically distributed databases. Algorithms focus on: a) discovering useful knowledge about the similarities and differences between the distributed databases, and b) construction of predictive models with high accuracy. Papers: J1, J4, A2, C3, C4, E1.
- *Reinforcement Learning*. Development of algorithms for efficient reinforcement learning, by utilizing transfer learning techniques. Papers: J17, C28, C36.
- *Applications*. Application of machine learning algorithms for the analysis of real-world data (biology, environment, music, software engineering, recommender systems), with the aim of bringing out their practical value. Papers: J5, J7, J9, J16, A1, C1, C12, C13, C23, C25, C27, C29, C40, C41, C42, C44, C45.

## 10. Publications

I have edited 1 volume of papers and published 2 theses and 89 papers (6 chapters in books, 24 journal articles, 1 encyclopedia lemma and 58 papers in international conference/workshop proceedings).

### Theses

- T1. Tsoumakas, G. (2000) "*A Knowledge Based System for Rock Fracture Classification in Deep Boreholes*", M.Sc. Thesis, Division of Informatics, University of Edinburgh, U.K., September 2000.
- T2. Tsoumakas, G. (2005) "*Machine Learning for the Fusion of Multiple, Distributed Intelligent Systems*", PhD Thesis (in Greek), Department of Informatics, Aristotle University of Thessaloniki, July 2005.

### Edited Volumes

- V1. Iliadis, L., Maglogiannis, I., Tsoumakas, G., Vlahavas, I., Bramer, M. (2009) "Artificial Intelligence Applications and Innovations: Proceedings of the 5th IFIP Conference on Artificial Intelligence Applications and Innovations (AIAI'2009)", April 23-25, 2009, Vol. 2962009, VIII, ISBN: 978-1-4419-0220-7.



## Book Chapters

- A1. Tsoumakas, G., Vlahavas, I. (2001) "Land Evaluation - An Artificial Intelligence Approach", in *Environmental Information Systems in Industry and Public Administration*, C. Rautenstrauch and S. Patig (eds.), Ch. 9, pp. 158-166, Idea Group Publishing.
- A2. Tsoumakas, G., Bassiliades, N., Vlahavas, I. (2004) "A Knowledge-based Web Information System for the Federation of Distributed Classifiers", in *Web Information Systems*, D. Taniar and W. Rahayu (eds.), Ch. 8, pp. 271-308, Idea Group Publishing.
- A3. Vrakas, D., Tsoumakas, G., Bassiliades, N., Vlahavas I. (2005) "Machine Learning for Adaptive Planning", in *Intelligent Techniques for Planning*, I. Vlahavas and D. Vrakas (eds.), Ch. 3, pp. 90-120, Idea Group Publishing.
- A4. Katakis, I., Tsoumakas, G., Vlahavas I. (2006) "Email Mining: Emerging Techniques for Email Management", in *Web Data Management Practices: Emerging Techniques and Technologies*, A. Vakali and G. Pallis (eds.), Ch. 10, pp 219-241, Idea Group Publishing.
- A5. Tsoumakas, G., Partalas, I., Vlahavas, I. (2009) "An Ensemble Pruning Primer", in *Applications of Supervised and Unsupervised Ensemble Methods*, G. Valentini and O. Okun (eds.), Studies in Computational Intelligence, Volume 245, Springer, pp 1-13.
- A6. Tsoumakas, G., Katakis, I., Vlahavas, I. (2010) "Mining Multi-Label Data", Part 6, Chapter 34, in *Data Mining and Knowledge Discovery Handbook*, L. Rokach and O. Maimon (eds.), Springer, pp. 667 – 285.

## Journal Articles

- J1. Tsoumakas, G., Angelis L., Vlahavas, I. (2004) "Clustering Classifiers for Knowledge Discovery from Physically Distributed Databases", *Data & Knowledge Engineering*, 49(3), pp. 223-242, Elsevier.
- J2. Vrakas, D., Tsoumakas, G., Bassiliades, N., Vlahavas, I. (2005) "HAP<sub>RC</sub>: An Automatically Configurable Planning System", *AI Communications*, 18(1), pp. 41-60, IOS Press.
- J3. Tsoumakas, G., Angelis, L., Vlahavas, I. (2005) "Selective Fusion of Heterogeneous Classifiers", *Intelligent Data Analysis*, 9(6), pp. 511-525, IOS Press.
- J4. Tsoumakas, G., Vlahavas, I. (2007) "A Scalable and Interoperable System for Classifier Sharing and Fusion", *Expert Systems with Applications*, 33(3), pp. 716-724, Elsevier.
- J5. Vrakas, D., Tsoumakas, G., Kokkoras, F., Bassiliades, N. Vlahavas, I., Anagnostopoulos, D. (2007) "PASER: A Curricula Synthesis System based on Automated Problem Solving", *International Journal on Teaching and Case Studies (IJTCS)*, 1(1/2), pp. 159-170, Inderscience.
- J6. Tsoumakas, G., Katakis, I. (2007) "Multi-Label Classification: An Overview", *International Journal of Data Warehousing and Mining*, 3(3), pp. 1-13, Idea Group Publishing.
- J7. Bibi, S., Tsoumakas, G., Vlahavas, I., Stamelos, I. (2008) "Regression via Classification applied on Software Defect Prediction", *Expert Systems with Applications*, 34(3), pp 2091-2101, Elsevier.
- J8. Katakis, I., Tsoumakas, G., Banos, E., Bassiliades, N., Vlahavas, I. (2008) "An Adaptive Personalized News Dissemination System", *Journal of Intelligent Information Systems*, 32(2), pp. 191-212, Springer.
- J9. Hatzikos, E.V., Tsoumakas, G., Tzanis, G., Bassiliades, N., Vlahavas, I. (2008) "An Empirical Study on Sea Water Quality Prediction", *Knowledge Based Systems*, 21(6), pp. 471-478, Elsevier.
- J10. Partalas, I., Tsoumakas, G., Hatzikos, E.V., Vlahavas, I. (2008) "Greedy Regression Ensemble Selection: Theory and an Application to Water Quality Prediction", *Information Sciences*, 178(20), pp. 3867-2879, Elsevier.
- J11. Partalas, I., Tsoumakas, G., Vlahavas, I. (2009) "Pruning an Ensemble of Classifiers via Reinforcement Learning", *Neurocomputing*, 72(7-9), pp. 1900-1909, Elsevier.
- J12. Katakis, I., Tsoumakas, G., Vlahavas, I. (2010) "Tracking Recurring Contexts using Ensemble Classifiers: An Application to Email Filtering", *Knowledge and Information Systems*, 22(3), pp. 371-391.
- J13. I. Partalas, G. Tsoumakas, I. Vlahavas, (2010) "An Ensemble Uncertainty Aware Measure for Directed Hill Climbing Ensemble Pruning", *Machine Learning*, 81(3), pp. pp. 257-282.
- J14. Tsoumakas, G., Katakis, I., Vlahavas, I. (2011) "Random k-Labelsets for Multi-Label Classification", *IEEE Transactions on Knowledge and Data Engineering*, 23(7), pp. 1079-1089, 2011

- J15. Tsoumakas, G., Spyromitros-Xioufis, E., Vilcek, J., Vlahavas, I. (2011) Mulan: A Java Library for Multi-Label Learning, *Journal of Machine Learning Research*, 12, pp. 2411-2414
- J16. Trohidis, K., Tsoumakas, G., Kalliris, G., Vlahavas, I. (2011) "Multi-label classification of music by emotion", *EURASIP Journal on Audio, Speech, and Music Processing* 2011, 2011:4, Published: 18 September 2011
- J17. Fachantidis, A., Partalas, I., Tsoumakas, G., Vlahavas, I. (2013) "Transferring Task Models in Reinforcement Learning Agents", *Neurocomputing*, 107, pp. 23-32
- J18. Spyromitros-Xioufis, E., Papadopoulos, S., Kompatsiaris, I., Tsoumakas, G., Vlahavas, I. (2014) "A Comprehensive Study over VLAD and Product Quantization in Large-scale Image Retrieval", *IEEE Transactions on Multimedia*, 16(6), pp.1713-1728
- J19. Markatopoulou, F., Tsoumakas, G., Vlahavas, I. (2015) "Dynamic Ensemble Pruning based on Multi-Label Classification", *Neurocomputing*, Elsevier, Volume 150, Part B, pp. 501-512
- J20. Spolaôr, N., Monard, M.C., Tsoumakas, G., Lee, H.D. (2015) A systematic review of multi-label feature selection and a new method based on label construction, *Neurocomputing*, Volume 180, pp. 3-15.
- J21. Kavakiotis, I., Triantafyllidis, A., Ntelidou, D., Alexandri, P., Megens, H.J., Crooijmans, R.P., Groenen, M.A., Tsoumakas, G., Vlahavas, I. (2015) TRES: Identification of Discriminatory and Informative SNPs from Population Genomic Data. *Journal of Heredity*, Volume 106(5):672-6.
- J22. Spyromitros-Xioufis, E., Tsoumakas, G., Groves, W., Vlahavas, I. (2016) Multi-Target Regression via Input Space Expansion: Treating Targets as Inputs, *Machine Learning*, Johannes Fürnkranz (Ed.), Springer, 2016.
- J23. Kavakiotis I., Xochelli A., Agathangelidis A., Tsoumakas G., Maglaveras N., Stamatopoulos K., Hadzidimitriou A., Vlahavas I., Chouvarda I. (2016). Integrating Multiple Immunogenetic Data Sources for Feature Extraction and Mining Mutation Patterns: The Case of Chronic Lymphocytic Leukemia Shared Mutations. *BMC Bioinformatics* 17(Suppl 5):173, pp. 375-387.
- J24. Papanikolaou Y., Foulds, J., Rubin, T., Tsoumakas G. (2017). Dense Distributions from Sparse Samples: Improved Gibbs Sampling Parameter Estimators for LDA. *Journal of Machine Learning Research* 18(62), pp. 1-58.

## Poster Presentations

- P1. Tsoumakas,G.; Spyromitros-Xioufis,E.;Vilcek,J.;Vlahavas,I. "Mulan: A Java Library for Multi-Label Learning", ICML 2010 Workshop on Machine Learning Open Source Software (MLOSS 2010), Demo/Poster Session, Haifa, Israel, 2010.

## Encyclopedia Lemmata

- E1. Tsoumakas, G., Vlahavas, I. (2008) "Distributed Data Mining", in *Encyclopedia of Data Warehousing and Mining – 2nd Edition*, J. Wang (ed.), pp. 709-715, Idea Group Reference.

## Papers in Conference and Workshop Proceedings

- C1. Tsoumakas, G., Vlahavas, I. (1999) "ISLE: An Intelligent System for Land Evaluation", *Proc. ACAI'99 Workshop on Intelligent Techniques for Spatio-Temporal Data Analysis in Environmental Applications*, pp. 26-32, Chania, Greece.
- C2. Tsoumakas, G., Vlahavas, I. (2001) "Fuzzy Meta-Learning: Preliminary Results", *Proc. UK Workshop on Computational Intelligence (UKCI 01)*, pp. 43-48, Edinburgh, September 2001.
- C3. Tsoumakas, G., Vlahavas, I. (2002) "Distributed Data Mining of Large Classifier Ensembles", *Proc. (Companion Volume) 2<sup>nd</sup> Hellenic Conference on AI (SETN 2002)*, pp. 249-255, Thessaloniki, April 2002.
- C4. Tsoumakas, G., Vlahavas, I. (2002) "Effective Stacking of Distributed Classifiers", *Proc. 15<sup>th</sup> European Conference on Artificial Intelligence (ECAI 2002)*, IOS Press, pp. 340-344, Lyon, France, July 2002.
- C5. Vrakas, D., Tsoumakas, G., Vlahavas, I. (2002) "Towards adaptive Heuristic Planning through Machine Learning", *Proc. 21st Workshop of the UK Planning and Scheduling Special Interest Group (PlanSIG 2002)*, pp 12-21, Delft, The Netherlands, November 2002.

- C6. Vrakas, D., Tsoumakas, G., Bassiliades, N., Vlahavas, I. (2003) "Learning Rules for Adaptive Planning", *Proc. 13<sup>th</sup> International Conference on Automated Planning and Scheduling (ICAPS 2003)*, pp. 82-91, Trento, Italy, June 2003.
- C7. Tsoumakas, G., Vrakas, D., Bassiliades, N., Vlahavas, I. (2004) "Using the k Nearest Problems for Adaptive Multicriteria Planning", *Proc. 3<sup>rd</sup> Hellenic Conference on Artificial Intelligence (SETN 2004)*, Springer-Verlag, LNAI 3025, pp. 132-141, Samos, Greece, May 2004.
- C8. Tsoumakas, G., Vrakas, D., Bassiliades, N., Vlahavas, I. (2004) "Lazy Adaptive Multicriteria Planning", *Proc. 16th European Conference on Artificial Intelligence (ECAI 2004)*, IOS Press, pp. 693-697, Valencia, Spain, August 2004.
- C9. Tsoumakas, G., Katakis, I., Vlahavas, I. (2004) "Effective Voting of Heterogeneous Classifiers", *Proc. 15th European Conference on Machine Learning (ECML 2004)*, Springer-Verlag, LNAI 3201, pp. 465-476, Pisa, Italy, September 2004.
- C10. Tsoumakas, G., Meditskos, G., Vrakas, D., Bassiliades, N., Vlahavas, I. (2004) "Web Services for Adaptive Planning", *Proc. ECAI 2004 Workshop on Planning and Scheduling: Bridging Theory to Practice*, Valencia, Spain, August, 2004. Also appeared in *Planning, Scheduling and Constraint Satisfaction: From Theory to Practice*, Volume 117, Frontiers in Artificial Intelligence and Applications, L. Castillo, D. Borrajo, M.A. Salido and A. Oddi, (Eds.), pp. 159-168, IOS Press, February 2005.
- C11. Katakis, I., Tsoumakas, G., Vlahavas, I. (2005) "On the Utility of Incremental Feature Selection for the Classification of Textual Data Streams", *Proc. 10<sup>th</sup> Panhellenic Conference on Informatics (PCI 2005)*, pp. 338-348, Volos, Greece, November 2005.
- C12. Diplaris, S., Tsoumakas, G., Mitkas, P., Vlahavas, I. (2005) "Protein Classification with Multiple Algorithms", *Proc. 10<sup>th</sup> Panhellenic Conference on Informatics (PCI 2005)*, pp. 448-456, Volos, Greece, November 2005.
- C13. Bibi, S., Tsoumakas, G., Stamelos, I., Vlahavas, I. (2006) "Software Defect Prediction Using Regression via Classification", *Proc. 4th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA '06)*, pp. 330-337, Dubai/Sharjah, March 8-11, 2006.
- C14. Partalas, I., Tsoumakas, G., Katakis, I., Vlahavas, I. (2006) "Ensemble Pruning using Reinforcement Learning", *Proc. 4<sup>th</sup> Hellenic Conference on Artificial Intelligence (SETN 2006)*, LNAI 3955, pp 301-310, Heraklion, Greece, May 18-20, 2006.
- C15. Tsoumakas, G., Katakis, I., Vlahavas, I. (2006) "A Review of Multi-Label Classification Methods", *Proc. 2<sup>nd</sup> ADBIS Workshop on Data Mining and Knowledge Discovery (ADMKD 2006)*, pp 99-109, Thessaloniki, Greece September 2006.
- C16. Katakis, I., Tsoumakas, G., Vlahavas, I. (2006) "Dynamic Feature Space and Incremental Feature Selection for the Classification of Textual Data Streams", *Proc. ECML/PKDD-2006 International Workshop on Knowledge Discovery from Data Streams (IWKDD-2006)*, pp 107-116, Berlin, Germany, September 2006.
- C17. Banos, E., Katakis, I., Bassiliades, N., Tsoumakas, G., Vlahavas, I. (2006) "PersoNews: A Personalized News Reader Enhanced by Machine Learning and Semantic Filtering", *Proc. 5th International Conference on Ontologies, DataBases, and Applications of Semantics (ODBASE 2006)*, Springer-Verlag, pp. 975-982, Montpellier, France, 2006.
- C18. Partalas, I., Hatzikos, E.V., Tsoumakas, G., Vlahavas, I. (2007) "Ensemble Selection for Water Quality Prediction", *Proc. 10th International Conference on Engineering Applications of Neural Networks (EANN 2007)*, Thessaloniki, Greece, 29-31 August 2007.
- C19. Tsoumakas, G., Vlahavas, I. (2007) "Random k-Labelsets: An Ensemble Method for Multilabel Classification", *Proc. 18th European Conference on Machine Learning (ECML 2007)*, Warsaw, Poland, 17-21 September 2007.
- C20. Partalas, I., Tsoumakas, G., Vlahavas, I. (2008) "Focused Ensemble Selection: A Diversity-Based Method for Greedy Ensemble Selection", *Proc. 18th European Conference on Artificial Intelligence (ECAI'08)*, Patras, Greece, 21-25 July 2008.
- C21. Katakis, I., Tsoumakas, G., Vlahavas, I. (2008) "An Ensemble of Classifiers for coping with Recurring Contexts in Data Streams", *Proc. 18th European Conference on Artificial Intelligence (ECAI'08)*, Patras, Greece, 21-25 July 2008.

- C22. Tsoumakas, G., Partalas, I., Vlahavas, I. (2008) "A Taxonomy and Short Review of Ensemble Selection", *Proc. ECAI 08 Workshop on Supervised and Unsupervised Ensemble Methods (SUEMA-08)*, 41-46, July 2008.
- C23. Trohidis, K., Tsoumakas, G., Kalliris, G., Vlahavas, I. (2008) "Multi-label Classification of Music into Emotions", *Proc. 9th International Conference on Music Information Retrieval (ISMIR 2008)*, Philadelphia, USA, 14-18 September 2008, pp 325-330.
- C24. Tsoumakas, G., Katakis, I., Vlahavas, I. (2008) "Effective and Efficient Multilabel Classification in Domains with Large Number of Labels", *Proc. ECML/PKDD 2008 Workshop on Mining Multidimensional Data (MMD'08)*, Antwerp, Belgium, 2008.
- C25. Katakis, I., Tsoumakas, G., Vlahavas, I. (2008) "Multilabel Text Classification for Automated Tag Suggestion", *Proc. ECML/PKDD 2008 Discovery Challenge Workshop*, Antwerp, 2008.
- C26. Spyromitros, E., Tsoumakas, G., Vlahavas, I. (2008) "An Empirical Study of Lazy Multilabel Classification Algorithms", *Proc. 5th Hellenic Conference on Artificial Intelligence (SETN 2008)*, Syros, Greece, 2-4 October 2008.
- C27. Katakis, I., Meditskos, G., Tsoumakas, G., Bassiliades, N., Vlahavas, I. (2009) "On the Combination of Textual and Semantic Descriptions for Automated Semantic Web Service Classification", *Proc. 5th IFIP Conference on Artificial Intelligence Applications & Innovations (AIAI 2009)*, Thessaloniki, Greece, 23-25 April 2009.
- C28. Partalas, I., Tsoumakas, G., Tzevanidis, K., Vlahavas, I. (2009) "Transferring Experience in Reinforcement Learning through Task Decomposition", *Proc. 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2009) – Volume 2*, Budapest, Hungary, 2009, pp. 1193-1194.
- C29. Dimou, A., Tsoumakas, G., Mezaris, V., Kompatsiaris, I., Vlahavas, I. (2009) "An Empirical Study Of Multi-Label Learning Methods For Video Annotation", *Proc. 7th International Workshop on Content-Based Multimedia Indexing (CBMI 2009)*, Chania, Crete, 3-5 June 2009.
- C30. Nasierding, G., Tsoumakas, G., Kouzani, A. (2009) "Clustering-Based Multi-Label Classification for Image Annotation and Retrieval", *Proc. 2009 IEEE International Conference on Systems, Man and Cybernetics*, October 11-14, 2009, San Antonio, Texas, USA.
- C31. Tsoumakas, G., Dimou, A., Spyromitros, E., Mezaris, V., Kompatsiaris, I., Vlahavas, I. (2009) Correlation-Based Pruning of Stacked Binary Relevance Models for Multi-Label Learning, *Proc. ECML/PKDD 2009 Workshop on Learning from Multi-Label Data*, Bled, Slovenia, pp. 101-116.
- C32. Tsoumakas, G., Loza Mencia, E., Katakis, I., Park, S.-H., Fürnkranz, J. (2009) "On the Combination of Two Decompositive Multi-Label Classification Methods", *Proc. ECML/PKDD 2009 Workshop on Preference Learning*, Bled, Slovenia, pp. 114-129.
- C33. Markatopoulou, F., Tsoumakas, G., Vlahavas, I. (2010) "Instance-Based Ensemble Pruning via Multi-Label Classification", *22nd International Conference on Tools with Artificial Intelligence*, Arras, France, 27-29 October 2010, IEEE.
- C34. M. Ioannou, G. Sakkas, G. Tsoumakas, I. Vlahavas, (2010) "Obtaining Bipartitions from Score Vectors for Multi-Label Classification", *Proc. 22nd International Conference on Tools with Artificial Intelligence*, Arras, France, 27-29 October 2010, IEEE.
- C35. Nasierding, G., Kouzani, A.Z., Tsoumakas, G. (2010) A Triple-Random Ensemble Classification Method for Mining Multi-label Data, *Proc. 2010 IEEE International Conference on Data Mining Workshops*, pp. 49-56, 2010
- C36. Fachantidis, A., Partalas, I., Tsoumakas, G., Vlahavas, I. (2011) "Transferring Models in Hybrid Reinforcement Learning Agents", *Proc. 12th Conference on Engineering Application of Neural Networks*, 2011.
- C37. Sechidis, K., Tsoumakas, G., Vlahavas, I. (2011) "On the Stratification of Multi-Label Data", *Proceedings of ECML PKDD 2011*, Athens, Greece, 2011.
- C38. Spyromitros-Xioufis, E., Tsoumakas, G., Vlahavas, I. (2011) "Multi-label Learning Approaches for Music Instrument Recognition", *Proc. 9th International Symposium on Methodologies for Intelligent Systems (ISMIS 2011)*, Warsaw, Poland, pp. 734-743.
- C39. Spyromitros-Xioufis, E., Spiliopoulou, M., Tsoumakas, G., Vlahavas, I. (2011) "Dealing with Concept Drift and Class Imbalance in Multi-label Stream Classification", *Proc. 22nd International Conference on Artificial Intelligence (IJCAI 2011)*, pp. 1583-1588, AAAI press, Barcelona, Spain.

- C40. Spyromitros-Xioufis, E., Stachtari, E., Tsoumakas, G., Vlahavas, I. (2011) A Hybrid Approach for Cold-start Recommendations of Videlectures, Proc. ECML PKDD 2011 Discovery Challenge Workshop, Athens, Greece.
- C41. Spyromitros-Xioufis, E., Sechidis, K., Tsoumakas, G., Vlahavas, I. (2011) MLKD's Participation at the CLEF 2011 Photo Annotation and Concept-Based Retrieval Tasks, Proceedings CLEF 2011.
- C42. Spyromitros-Xioufis, E., Papadopoulos, S., Kompatsiaris, I., Tsoumakas, G., Vlahavas, I. (2012) An empirical study on the combination of surf features with VLAD vectors for image search, Proceedings 13th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), IEEE.
- C43. Spolaor, N., Tsoumakas, G. (2013) Evaluating Feature Selection Methods for Multi-Label Text Classification, BioASQ Workshop, Valencia, Spain, September 27, 2013
- C44. Tsoumakas, G., Laliotis, M., Markantonatos, N., Vlahavas, I. (2013) Large-Scale Semantic Indexing of Biomedical Publications at BioASQ, BioASQ Workshop, Valencia, Spain, September 27, 2013, 2013.
- C45. F. Briggs, Yonghong Huang, R. Raich, K. Eftaxias, Zhong Lei, W. Cukierski, S. Hadley, A. Hadley, M. Betts, X. Fern, J. Irvine, L. Neal, A. Thomas, G. Fodor, G. Tsoumakas, Hong Wei Ng, Thi Ngoc Tho Nguyen, H. Huttunen, P. Ruusuvaori, T. Manninen, A. Diment, T. Virtanen, J. Marzat, J. Defretin, D. Callender, C. Hurlburt, K. Larrey, M. Milakov, "The 9th annual MLSP competition: New methods for acoustic classification of multiple simultaneous bird species in a noisy environment", Proc. 2013 IEEE International Workshop on Machine Learning for Signal Processing (MLSP), IEEE, 2013.
- C46. I. Kavakiotis, A. Triantafyllidis, G. Tsoumakas, I. Vlahavas, "Feature Evaluation Metrics for Population Genomic Data", Proceedings of 8th Hellenic Conference on Artificial Intelligence (SETN 2014), A. Likas, K. Blekas and D. Kalles (Eds.), Springer, Artificial intelligence: Methods and Applications, LNCS, 8445, pp. 436-441, Ioannina, Greece, May 15-17, 2014, 2014.
- C47. Tsoumakas, G., Spyromitros-Xioufis, E., Vrekou, A., Vlahavas, I. (2014) Multi-Target Regression via Random Linear Target Combinations, Proceedings of ECML PKDD 2014, Nancy, France, pp. 225-240.
- C48. Papanikolaou, Y., Dimitriadis, D., Tsoumakas, G., Laliotis, M., Markantonatos, N., Vlahavas, I. (2014) Ensemble Approaches for Large-Scale Multi-Label Classification and Question Answering in Biomedicine. CLEF (Working Notes), pp. 1348-1360.
- C49. Arvanitidis, A., Serafi, A., Vakali, A., Tsoumakas, G. (2014) Branty: A Social Media Ranking Tool For Brands, Proceedings of ECML PKDD 2014 (Demo Track), Springer, Nancy, France, pp 432-435.
- C50. Tsoumakas, G., Papadopoulos, A., Qian, W., Vologianidis, S., D'yakonov, A., Puurula, A., Read, J., Švec, J., Semenov, S. (2014) WISE 2014 challenge: Multi-label classification of print media articles to topics, Proceedings of the 2014 Web Information Systems Engineering Conference, pp. 541-548.
- C51. Spolaor, N., Monard, M.C., Tsoumakas, G., Lee, H.D. (2014) Label construction for multi-label feature selection, Proceedings of the 2014 Brazilian Conference on Intelligent Systems (BRACIS), pp. 247-252.
- C52. Papagiannopoulou, C., Tsoumakas, G., Tsamardinos, I. (2015) Discovering and Exploiting Deterministic Label Relationships in Multi-Label Learning, Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 915-924.
- C53. Papagiannopoulou, E., Tsoumakas, G., Bassiliades, N. (2015) On Discovering Relationships in Multi-Label Learning via Linked Open Data, 4th Workshop on Knowledge Discovery and Data Mining Meets Linked Open Data (Know@LOD 2015), Johanna Völker, Heiko Paulheim, Jens Lehmann, Vojtech Svatek (Ed.), CEUR Workshop Proceedings, Vol-1365.
- C54. Papanikolaou, Y., Tsoumakas, G., Laliotis, M., Markantonatos, N., Vlahavas, I. (2015) AUTH-Atypon at BioASQ 3: Large-Scale Semantic Indexing in Biomedicine. Working Notes of CLEF 2015 - Conference and Labs of the Evaluation Forum, CEUR Workshop Proceedings, Vol-1391.
- C55. Samaras, P., Fachantidis, A., Tsoumakas, G., Vlahavas, I. (2015) A prediction model of passenger demand using AVL and APC data from a bus fleet, Proceedings of the 19th Panhellenic Conference on Informatics, 129-134.
- C56. Fachantidis A., Tsiaras A., Tsoumakas G. & Vlahavas I. (2016). Segmento: An R-based Visualization-rich System for Customer Segmentation and Targeting, Proc. 9th Hellenic Conference on Artificial Intelligence (SETN 2016), Thessaloniki, Greece, 18-21 May 2016.
- C57. Papagiannopoulou, E, Papanikolaou, Y., Dimitriadis, D., Lagopoulos, A., Tsoumakas, G., Laliotis, M., Markantonatos, N., Vlahavas I. (2016) "Large-Scale Semantic Indexing and Question Answering in Biomedicine", In Proceedings of the BioASQ 2016 Workshop.

- C58. Lagopoulos, A., Fachantidis, A., Tsoumakas, G. (2017) Multi-Label Modality Classification for Figures in Biomedical Literature, Proc. 30th IEEE International Symposium on Computer-Based Medical Systems.
- C59. Anagnostou, A., Lagopoulos, A., Tsoumakas, G., Vlahavas, I. (2017) Combining Inter-Review Learning-to-Rank and Intra-Review Incremental Training for Title and Abstract Screening in Systematic Reviews, eHealth Lab of the 8th Conference and Labs of the Evaluation Forum (CLEF).

Papers C41, C45 and C50 were accepted without peer-review.

## 11. Citations

According to Google Scholar my papers have attracted more than 7000 citations and I have an h-index of 34.

## 12. Other Information

### Acceptance ratio of selected conferences

Paper	Conference	Accepted	Ratio	Perc.
C19	18 <sup>th</sup> European Conference on Machine Learning (ECML 2007)	54/595	1:11	9 %
C9	15 <sup>th</sup> European Conference on Machine Learning (ECML 2004)	45/312	1:6.9	15 %
C39	22 <sup>nd</sup> International Joint Conference on Artificial Intelligence (IJCAI 2011)	227/1325	1:5.8	17 %
C52	KDD 2015	160/819	1:5.1	20 %
C37	ECML PKDD 2011	121/599	1:5	20 %
C20	18 <sup>th</sup> European Conference on Artificial Intelligence (ECAI 2008)	121/561	1:4.6	22 %
C48	ECML PKDD 2014	115/483	1:4.2	24 %
C4	15 <sup>th</sup> European Conference on Artificial Intelligence (ECAI 2002)	140/505	1:3.6	28 %
C8	16 <sup>th</sup> European Conference on Artificial Intelligence (ECAI 2004)	181/652	1:3.6	28 %