**Project no.: 619209**

**Project full title: Analysis of Massive Data STreams**

**Project Acronym: AMIDST**

**Deliverable no.: D2.1**

**Title of the deliverable: The AMIDST modelling framework – Initial draft report**

|  |  |
| --- | --- |
| **Contractual Date of Delivery to the CEC:** | **30.09.2014** |
| **Actual Date of Delivery to the CEC:** | **30.09.2014** |
| **Organisation name of lead contractor for this deliverable:** | **AAU** |
| **Author(s):** | **Hanen Borchani, Antonio Fernández, Odd E. Gundersen, Sigve Hovda, Helge Langseth, Anders L. Madsen, Ana M. Martínez, Ramón Martínez, Andrés Masegosa, Thomas D. Nielsen, Antonio Salmerón, Frode Sørmo, Galia Weidl** |
| **Participants(s):** | **P01, P02, P03, P04, P05, P06, P07** |
| **Work package contributing to the deliverable:** | **WP2** |
| **Nature:** | **R** |
| **Version:** | **1.0** |
| **Total number of pages:** | **68** |
| **Start date of project**: | **1st January 2014 Duration: 36 month** |

|  |  |  |
| --- | --- | --- |
| **Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)** | | |
| **Dissemination Level** | | |
| **PU** | Public | **X** |
| **PP** | Restricted to other programme participants (including the Commission Services) |  |
| **RE** | Restricted to a group specified by the consortium (including the Commission Services) |  |
| **CO** | Confidential, only for members of the consortium (including the Commission Services) |  |

**Abstract:**

|  |
| --- |
| In this document, we explore aspects of the different distribution families exhibited by the data for the three use-case providers. Based on these findings and the expert knowledge provided, we present preliminary use-case-tailored model classes that address the application scenarios identified in Deliverable 1.2. Building on the commonalties of these specific model classes, we introduce the AMIDST modelling framework, which defines a probabilistic model class intended to support data-driven learning and integration of potential domain knowledge. |

**Keyword list:** AMIDST modelling framework, application scenarios, preliminary models, data analysis, Bayesian networks, dynamic Bayesian networks, conditional linear Gaussian models.