02 August 2021

IFC workshop on "Data science in central banking"

Part 1:

19-22 October 2021, Bank of Italy, virtual event on

"Data Science in Central Banking: Machine learning applications"

Part 2:

14-17 February 2022, BIS Basel, Switzerland, hybrid (virtual/physical TBC) event on

"Data Science in Central Banking: Applications and tools"

Organised by the IFC at the BIS with the active support of the Bank of Italy, the European Central Bank and the South African Reserve Bank, this workshop will bring together central banks, international organisations, national statistical offices and other interested stakeholders to share knowledge on emerging trends in data science, data engineering and information technologies with a broad audience of practitioners and technicians. We will look in depth at the state of adoption of data analytics and business intelligence techniques along with data transformation and big data ecosystems in participants' organisations. This event is intended to foster exchange, collaboration and understanding on the related interdisciplinary practices, use cases, and technologies and will also cover important topics such as data governance and data protection.

The first part of the workshop will be hosted by the Bank of Italy (BoI) on 19-22 October 2021 and the second part by the Bank for International Settlements (BIS) on 14-17 February 2022. The first part would focus on "Data Science in Central Banking: Machine learning applications" whereas the second part would be on "Data Science in Central Banking: Applications and tools".

Given the pending Covid-related uncertainties, the October 2021 event will be held virtually by the BoI whereas the event in February 2022 will be combined ie hybrid (virtual/physical) at the BIS in Basel. On each workshop day a maximum of three hours presentations are foreseen, in order to facilitate global participation.

The two workshops will focus on two main topics:

- Sharing of experience on tools, technologies, platforms;
- Illustration of use cases developed in participants' organisations.

Areas of interest include in particular (see detailed annex):

- Data science and engineering (eg machine learning)
- Big data technologies and use cases
- Big data wrangling and processing
- Data visualisation and business intelligence
- Data governance / security
- Platforms for open banking / regulatory compliance
- Generic statistical business process modelling

Contact

- IFC Secretariat: Rafael.Schmidt@bis.org; Rukmani.Vaithianathan@bis.org; IFC.secretariat@bis.org
- Bank of Italy: giuseppe.bruno@bancaditalia.it; Juri.Marcucci@bancaditalia.it

Important dates

- Abstract submission: 15 January 2021
- Acceptance notification: 31 January 2021
- Presentation submission: 30 September 2021 for workshop part 1 / 31 January 2022 for workshop part 2
- Paper submission: 12 October 2021 for workshop part 1 / 7 February 2022 for workshop part 2

Annex

Topics of interest	Tools (not limited to this list)
Big data technologies and platforms	Hadoop Ecosystem, Spark
Microservices and containerised platforms	Openshift
Adoption of deep learning tools	Keras, TensorFlow, PyTorch, MxNet
Deep learning for time series forecast	Generative adversarial networks, variational inference (variational autoencoders)
Platforms for regulatory compliance	LegalRuleML, Rosetta
Adopting new organisational structure in financial services with open banking	OpenAPI
Managing and enhancing existing processes with evolving SDMX standards	SDMX tools
Adoption of VTL in statistical data processing	VTL
Data Lab and interactive data access platforms	Jupyter, Dataiku, RStudio
Automated feature engineering	Featuretools
Natural language processing	Open vs commercial products; Word2vec, Glove, fastText and StarSpace
Text classification: smart tagging / labelling of documents	Refinitiv's Intelligent Tagging, Python's Spacy
Recommendation systems in document search	Case Recommender, Algorithmic frameworks for content and collaborative filtering
Document search using context-based ranking systems	Context sensitive ranking algorithms
Knowledge base using graphs	Google's Knowledge Graph
Visualisation tools	Tableau, Power BI, Plotly, Dash, D3.js
Strategies for collaboration and publishing documents with interactive graphs	Plotly, D3.js, RShiny
DevOps: automating the data processing and ML pipelines	Adoption of scheduler, orchestrator, version control, CI/CD, deployment tools, performance monitoring and alerting in data science platforms.
DataOps: metadata management practices while developing ML pipelines	Metadata tools and technologies supporting data lineage, data provenance and data catalogue.