Random Feature Mapping for Optical Computing

Man Fung Ho[1], Egor Manuylovich[2], Sergei Turitsyn[2], **David Saad**[1]
[1]Aston University, Birmingham, United Kingdom

Machine learning and AI technologies come with a high and unsustainable cost, both in terms of energy and computing power, leading to the search for specialised alternative computing paradigms. Recent developments in optical computing have demonstrated its advantages over traditional computing methods, such as higher efficiency due to its inherent parallelism and lower energy cost. Most current efforts have focused on deep learning models, reservoir computing or extreme learning machines, which are not well understood analytically. We present a new approach, rooted in the non-linear mapping of data to a higher dimensional feature space, where various computing tasks can be easily carried out and which has solid mathematical foundations. Moreover, such a mapping can be implemented using optical devices. Regression, classification and time series prediction can then be carried out using linear operators in the high-dimensional space, which can also be executed using an optical implementation of matrix-vector multiplication that are rapidly becoming more efficient.

Our approach is based on solid mathematical foundations, showing that difficult tasks, primarily in classification and regression, can become much easier to solve after a nonlinear mapping to a high-dimensional space. This also forms the foundations of Support Vector Machines (SVM) which obviates the mapping of data through the use of kernels. When it comes to optical computing, kernels become difficult to implement, while the mapping to high-dimensional spaces can be done efficiently using simple, fast and commercially available tools such as the Mach-Zehnder Modulator. The mapping method is based on the following steps: 1) Sampling a random vector from a given distribution which correspond to a specific kernel. 2) Then, a vector-matrix multiplication between the data and sampled vectors, is implemented optically using a Spatial Light Modulator (see figure). 3) The products are used to generate the vectors to be mapped via simple trigonometric functions, that can be implemented using optical devices (Mach-Zehnder Modulator), transforming the data to a high-dimensional space. 4) A vector-matrix multiplication followed by a linear operation is used to classify data or obtain the regression value. Moreover, we exploit a method, termed Random Feature Mapping, to obtain the appropriate random mappings per task, depending on the corresponding kernel. We then carry out numerical experiments for specific regression, classification and time series prediction tasks. The method was tested both in numerical simulations and on emulated optical devices, showing excellent results of similar level to the performance of SVM.