The Catalyst Project, funded by the Higher Education Funding Council for England (HEFCE) and monitored by the Office for Students (OfS) is a partnership between the University of Essex, Essex County Council and Suffolk County Council.

Using cross-disciplinary expertise in data analytics, big data and evaluation, the University of Essex is supporting both County Councils to assess risks for vulnerable members of the community and provide evaluation techniques to fully understand the impact of Council initiatives.

Catalyst Risk Stratification Case Study

School Readiness in Essex

The Catalyst Project Risk Stratification team, based at the University of Essex, used their expertise in predictive analytics to support a pilot initiative run by the Essex Data (ED) programme. The team applied machine learning methods to help predict levels of *School Readiness* of children from a specific community in Essex. The objective was to use data to predict the level of risk of not being school ready. Understanding in advance when children are not school ready enables targeted interventions that support families to prepare their children to be socially, physically and intellectually ready to start attending school aged five.

The study predicted 511 households that were at risk in the pilot area, 280 of those households were not known to public services. Consequently, this community, together with local authorities, are now co-designing services to give their children the skills they need to thrive at school. The initiative aims to improve children's longer term outcomes into adulthood by allowing them the best start in life and thereby reducing the risk of future service demand.

> "It is nice to have a reputable institution clearly demonstrate their own approach to verify our understanding of predictive analytics. I feel much more comfortable with the fundamental application of this sort of methodology – it has enabled us to optimise our data science approaches."
>
> **Stephen Simpkin, Data Science Fellow, Essex County Council**

The Essex Data (ED) programme

The Essex Data (ED) programme was funded by a grant awarded to Essex Partners. The ED programme enables public sector organisations across Essex to target support to address some of the most challenging local issues through smarter use of existing data and the adoption of advanced analytical methodologies. The ED programme facilitates the safe sharing and matching of multiple data sets, which can be used together to better understand the needs of vulnerable people in the county and to drive evidence led decision-making. A number of pilot initiatives focused on pressing issues have been implemented over the last two years to test the feasibility of the ED programme. One such pilot initiative focused on using data science to predict and improve *School Readiness* in Essex.

School Readiness is a term that describes a child's ability to engage in and benefit from early learning experiences. It is measured by a set of criteria relating to a child's personal and intellectual development before the age of five. This involves assessing cognitive abilities, social interaction, and attention span, amongst other factors. Evidence shows that children who are not ready for school and who do not meet the key developmental milestones experience additional challenges compared to their peers. Therefore helping children to be prepared for starting school will help them to achieve better outcomes, both in school and in their lives ahead.

Building the predictive model

The Catalyst Project Risk Stratification team worked alongside Essex Partners on the *School Readiness* pilot initiative by applying data analysis and advanced machine learning techniques to anonymised data accessed by the ED programme. The aim was to build and test a machine learning approach to assist Essex Partners to better understand how it could target services to support the community and prevent problems from arising. Data was collected from partner organisations, including information on attainment, demographics, crime, housing, benefits, income, and health.

Building the predictive model involved:

- preparation of a literature review to gain a clear understanding of existing research and methodologies

- exploration of the data and creation of descriptive statistics by Catalyst Project analysts to enable better understanding of the variables

- creation of a series of explanatory models to establish the most important variables in the dataset

- application of machine learning methods 'lasso' and 'random forest' to create models to identify children at risk of not being school ready

Catalyst Project outputs

The model built by the Catalyst Project Risk Stratification team has been used by Essex Partners to further develop the ED programme, which is part of the county's 'Future of Essex' strategy. The Catalyst Project has contributed academic expertise to the integration of machine learning techniques in the Council's decision making processes. Commissioners have been able to use insight from the ED programme to apply for £3.35 million funding for interventions, to inform work with Ofsted, and to form a 'New Generations' community group consisting of parents, head teachers and charity representatives, which has implemented a new nursery and family boot camp in the pilot area.

In summary, the Catalyst Project provided:

- an overview of the global literature on factors that can influence *School Readiness* and interventions that can help children overcome barriers

- a machine learning model with high predictive ability

- the code and methodology necessary to replicate the analysis with additional data (this was loaded into the Omniscope platform ready for use by Essex County Council analysts)

- a report identifying local area 'hotspots' including important risk and protective factors

- knowledge transfer in the form of training Council analysts to use machine learning algorithms in the programming language 'R'

Contact

E catalyst@essex.ac.uk

T +44 (0) 1206 872057

🐦 @HEFCE_Catalyst

https://www.essex.ac.uk/research-project/catalyst

Catalyst is funded by:

HIGHER EDUCATION *hefce* FUNDING COUNCIL FOR ENGLAND

University of Essex in partnership with:

Essex County Council

ESSEX PARTNERS