



## Scoring Decentralized Finance Protocols

APPROACH PAPER V1.0

## **1. Abstract:**

Decentralized Finance (DeFi) has gained significant traction in recent years. We believe that, in the future Decentralized Finance holds promise for it has the potential to democratize finance like never before; its potential is premised on its ability to comply with rules and regulations of the sovereign and its ability to continuously demonstrate that it does not intend to defraud the community that supports and nurtures it.

BetygFi is a comprehensive machine learning-based approach to assess DeFi protocols. The goal is to evolve a robust, real time, neutral, self-learning approach towards analyzing DeFi's.

The model underpinning Betygfi currently analyses technical risk, centralization risk, financial/market risk, and userbase quality risks. The model currently leverages a combination of supervised and unsupervised learning techniques to enable classification of DeFi protocols/projects.

## **2. Data Collection and Processing**

At the core of every machine learning model are dependable data sources/data sets, an efficient data pipeline, and dependable data normalization techniques. Solvendo, through its work in traditional finance understands the significance of a robust ETL (extraction, transformation and loading) methodology.

Below is a recital of the Datasets and Features that we collect, real-time, to train and evaluate our model. We currently compile data on the following features:

### **2.1 Technical Risk**

- Time on Mainnet
- Number of Critical Vulnerabilities
- Public Audit
- Recent Audit
- Open Source
- Byte Source code verified
- Number of Engineers working
- Number of code commits
- Bounty Program

### **2.2 Centralization Risk**

- Protocol Administration
- Private Key Model and arrangement
- Oracles
- Governance Token Concentration

### **2.3 Financial/Market Risk**

- Collateralization Ratio
- 30-day Exponential Moving Average (EMA)
- Liquidity 30-day EMA
- Total Value Locked (TVL) 30-day EMA
- Volume Weighted Average Price (VWAP)
- Relative Strength Index (RSI)
- Asset Breakdown

### **2.4 Userbase Quality**

- Repayment
- Active Loans
- Asset Quality
- Asset Liquidation Threshold

## **3. Model Training Methodology**

Our approach combines supervised and unsupervised machine learning techniques to derive a comprehensive scoring system for DeFi protocols. With reference to the ETL process mentioned above, currently we first preprocess data by normalizing features; then deal with the missing values and encode categorical variables. We thereafter perform the following steps:

### **3.1 Unsupervised Learning**

- Implement clustering algorithms (e.g., K-means, DBSCAN) to identify patterns and group DeFi protocols based on their inherent risk characteristics.
- Analyze the resulting clusters to derive insights into the risk profiles of different DeFi protocols.

### **3.2 Supervised Learning**

- Train classification models (e.g., Random Forest, Support Vector Machine, Neural Networks) on a labeled dataset, where each DeFi protocol is assigned a risk score based on expert opinion or historical performance.
- Evaluate the performance of the classifiers using metrics such as accuracy, precision, recall, and F1 score.
- Optimize the best-performing model using techniques such as hyperparameter tuning and feature selection.

#### 4. Model Evaluation:

After training and optimizing the model, we evaluate the model's performance on a test dataset. We use a myriad of approaches including comparing the model's generated scores with actual performance. We also analyze feature importance to understand the key factors contributing to the risk profiles of DeFi protocols.

#### 5. Scoring:

We truly believe that all the advances in machine learning including natural language processing (NLP) through foundation models, can best be utilized to navigate real world problems - by designing solutions that are binary but explainable. This approach enables users to understand and consume the output easily. This belief is at the core of Solvendo's thesis and existence and is also reflected in the approach at BetygFi.

The Solvendo score that underpins the efforts at BetygFi is intended to enable users to understand at a glance, the risk reward profile of DeFi protocols to further enable informed decision making. As indicated above currently the score comprises an analysis of DeFi's technical, centralization, financial/market, and userbase quality risks.

#### 6. Future:

The model's evolution will be reflected in the accuracy, predictive performance and therefore usefulness of the score. We have a broad outline of how we plan to train and evolve the model for the benefit of users and their better understanding of Defi protocols/projects.

Keeping with our intention to continuously improve the accuracy and efficiency of the model, our ongoing efforts include as under:

**6.1 Feature Set Expansion:** Additional features will continuously be included to enable a comprehensive view of the risks associated with DeFi protocols. In our opinion, incorporating data on protocol-specific features or emerging risks would enhance the model's performance.

**6.2 Real-Time Risk Assessment:** It is our continuous pursuit to make our ETL process real time; to enable us to integrate real time data into the model, thus to enable dynamic real time risk assessment for DeFi protocols. We believe that such real time risk assessment would be a game changer in risk assessment and management.

**6.3 Better Machine Learning Techniques:** We continue to explore advancements in machine learning. The space is fast evolving and is at the moment witnessing a gold rush. As a consequence of these efforts, we expect promising first and second layer approaches and use cases. For example, better deep learning or ensemble methods could potentially improve the predictive performance of the model. Implementing recurrent neural networks (RNNs) or

transformers to analyze time-series data or using stacking techniques to combine the predictions of multiple models could yield better results.

**6.4 Explainable Results:** Notwithstanding all of the benefits of machine learning, we anticipate black box models will face significant challenges, as results and use cases from such models become more mainstream. It is accepted that increase in complexity of the model makes it more challenging to interpret, complexity also makes it difficult to understand the logic behind the model's decisions. Implementing explainable AI techniques, such as Shapley Additive Explanations (SHAP) values or Local Interpretable Model-agnostic Explanations (LIME), usually help stakeholders better understand the underlying reasons behind the model's risk assessments. It is our constant endeavor to deploy the best available tools to enable us to best explain our efforts and results.

**6.5 Cross-Chain and Layer-2 Protocols:** As the DeFi ecosystem continues to evolve, we expect more cross-chain and layer-2 protocols to emerge. Incorporating these new technologies into our risk assessment framework will be crucial to enable us to continue to provide a comprehensive and up-to-date understanding of the DeFi landscape and associated risks.

Our approach is an attempt to explain transparently our value proposition and efforts. We hope to continuously strive to enable actionable insights for stakeholders in the decentralized financial ecosystem.