**Слайд 1.**

Добрый день. Меня зовут Фёдор. Я представляю работу на тему «Разработка и использование моделей классификации в кредитном скоринге (на примере ОАО «Белинвестбанк»)».

**Слайд 2.**

Сначала обсудим задачу, которая стала объектом данной работы;

Затем перейдем к идентификационной форме модели, использованной в этой работе;

Потом представим использованный набор данных и комплекс преобразований над ним для подготовки к моделированию;

Завершают мое выступления тонкости связанные с получением модели и её валидацией.

**Слайд 3.**

В Банк ежедневно поступает множество заявок **(анимация)** на выдачу кредита. Требуется проводить данных предоставленных клиентом для принятия решения о выдаче или удержании кредита. Когда речь касается физических лиц, а, на сегодняшний день, уже и малого бизнеса – объемы поступающих заявок настолько велики, что рассмотреть в ручную каждую заявку не представляется возможным. Для обработки поступающих заявок создается кредитный конвейер **(анимация)** – механизм который позволяет обрабатывать заявки автоматически. Сердцем которого становиться скоринговая модель **(анимация)** – механизм который позволяет автоматически принимать решение **(анимация)** о выдаче или удержании кредита для каждого потенциального кредитополучателя на основе наблюденной ранее статистики.

Наша задача заключалась в формировании такой модели. Понятно, что поставленная задача сводиться к задаче классификации. И формально **(анимация)** может быть описана так: сформировать некоторое правило R которое на основании входных данных X выдает бинарный оклик кодирующий выдачу или удержание кредита.

**Слайд 4.**

Изначально для решения такой задачи, на данных, которые мы представим далее, мы использовали классический алгоритм логистической регресии. Тут, сверху представлена идентификационная форма такой модели. Для нашей модели качества такого классификатора было недостаточно. Заметим, что это получается линейный классификатор, покажем это на примере **(анимация)** тут слева два класса точек расположились по двум переменным, если на таких данных сроить логистическую регрессию получатся предсказания верностей как отложены на втором рисунке. Выбрав некоторую вероятность после которой мы считаем клиента склонным к дефолту у нас получается отложить линию уровня **(анимация)**, которая и станет дискриминирующей в данном случае. Как видно, даже в такой простой задаче модель допускает ошибки.

Наша идея заключалась в том, чтобы усложнить функцию под логитом, чтобы «сломать» эту прямую в нужном месте **(анимация)** – перейти к кусочно линейному классификатору и, за счет этого, добиться лучших предикативных свойств модели.

**Слайд 5**

Для построения модели с описанными свойствами мы использовали идеи лежащие в нейронных сетях. Формально идентификационную форму такой модели удобно представлять в виде графа где дуга для каждого рассмотренного наблюдения передает некоторое число, а каждый узел представляет собой некоторое преобразование данных.

В каждом из преобразований присутствует сумматорная функция, которая представляет собой линейную комбинацию активаций предыдущего слоя – на графе это входящие в каждый нейрон дуги.