



AgroCode Hack 2022: Эффективные протоколы лечения коров

Команда: FEDOT HACK TEAM
Капитан: Стебеньков Андрей

AgroCode Hack
Пит-стоп #3
18.09.2022

СДЕЛАЛИ ЗА ХАКАТОН:

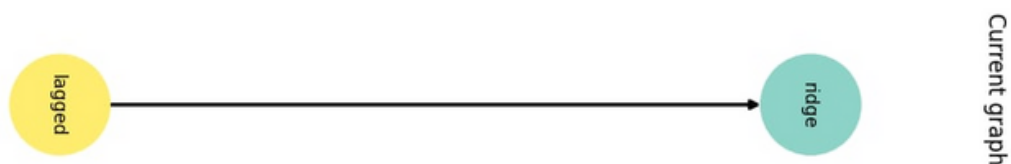
- Заполнение таблицы "Статистика" прямым методом (парсинг дампа);
- Анализ данных:
 - Проверили гипотезы зависимостей;
 - Нахождение вероятностей выбора удачного протокола лечения;
 - Нахождение резистентности коров к протоколу лечения;
 - Гипотеза зависимости Кетоза от сухостоя.
- Выявление связей через Байесовские сети:
 - Исправление предыдущего по Событиям;
 - Создания графа по переменным, которые влияют на успешность протокола.
- Генеалогический граф по коровам и болезнями

Заполнение таблицы "Статистика"

СДЕЛАЛИ:

- Получилось при помощи обработки таблицы Дамп заполнить все данные в таблице Статистика;
- Обучили модели для предсказания и заполнения статистик. Результат печальный
Вывод: Это не работает)

Готовы сверить заполненные данные
(файл: statistics_output.xlsx) с реальными!



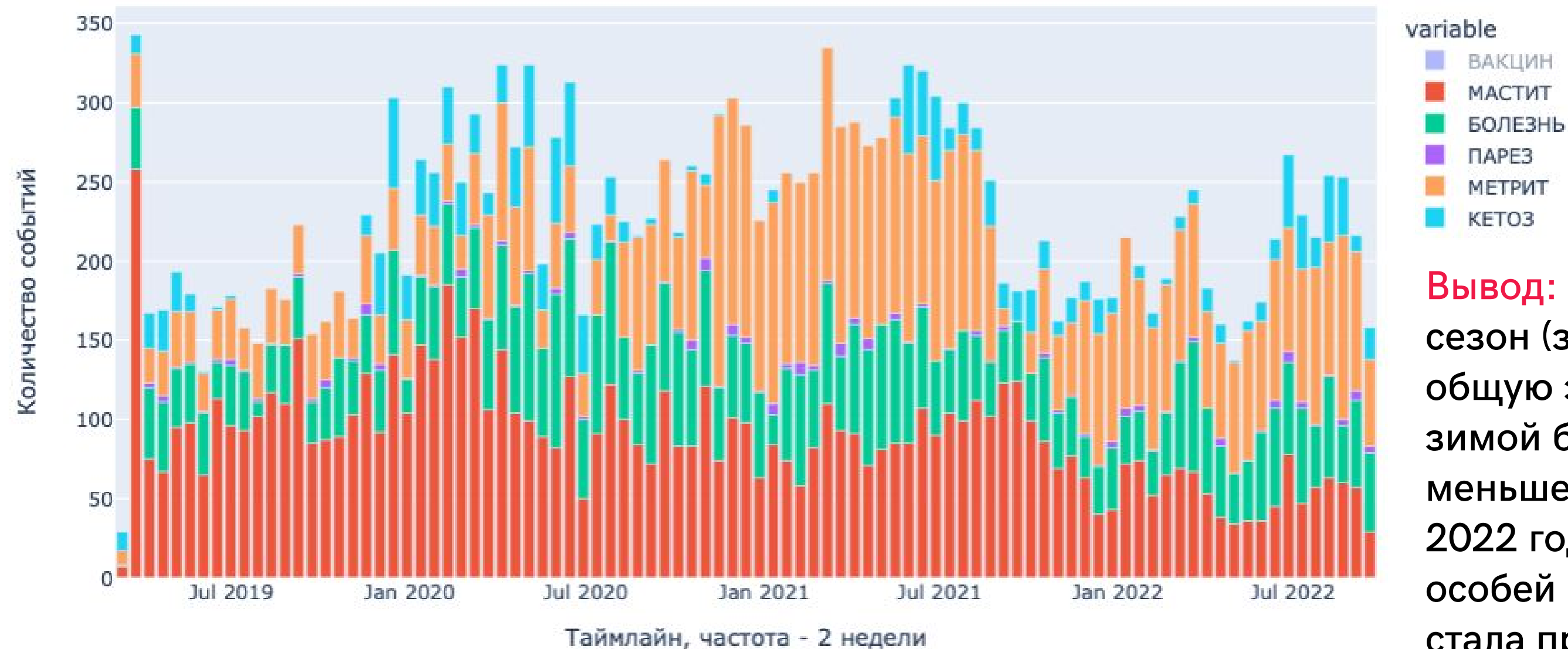
Полученная модель для предсказания
случаев за период

		Index	3.2022	4.2022	5.2022	6.2022	7.2022	8.2022	9.2022	10.2022	11.2022
0	Мастит случаев за период		126	84	79	132	132	130	46	0	0
1	Мастит голов за период		122	82	75	128	122	125	45	0	0
2	Передняя левая		30	18	27	33	42	41	18	0	0
3	Передняя правая		34	23	27	32	44	38	10	0	0
4	Задняя левая		58	43	38	53	54	61	26	0	0
5	Задняя правая		44	24	31	45	58	47	24	0	0
6	МВ		1	3	1	7	1	2	0	0	0
7	КМ5		1	1	22	16	6	6	10	0	0
8	КМ7		0	0	0	0	0	2	0	0	0
9	КМ1		41	20	26	46	49	54	15	0	0

Часть таблицы восстановленных данных

Анализ данных: Изучение болезней и событий

Заболеваемость (высота столба = сумма выявленных заболеваний за период)



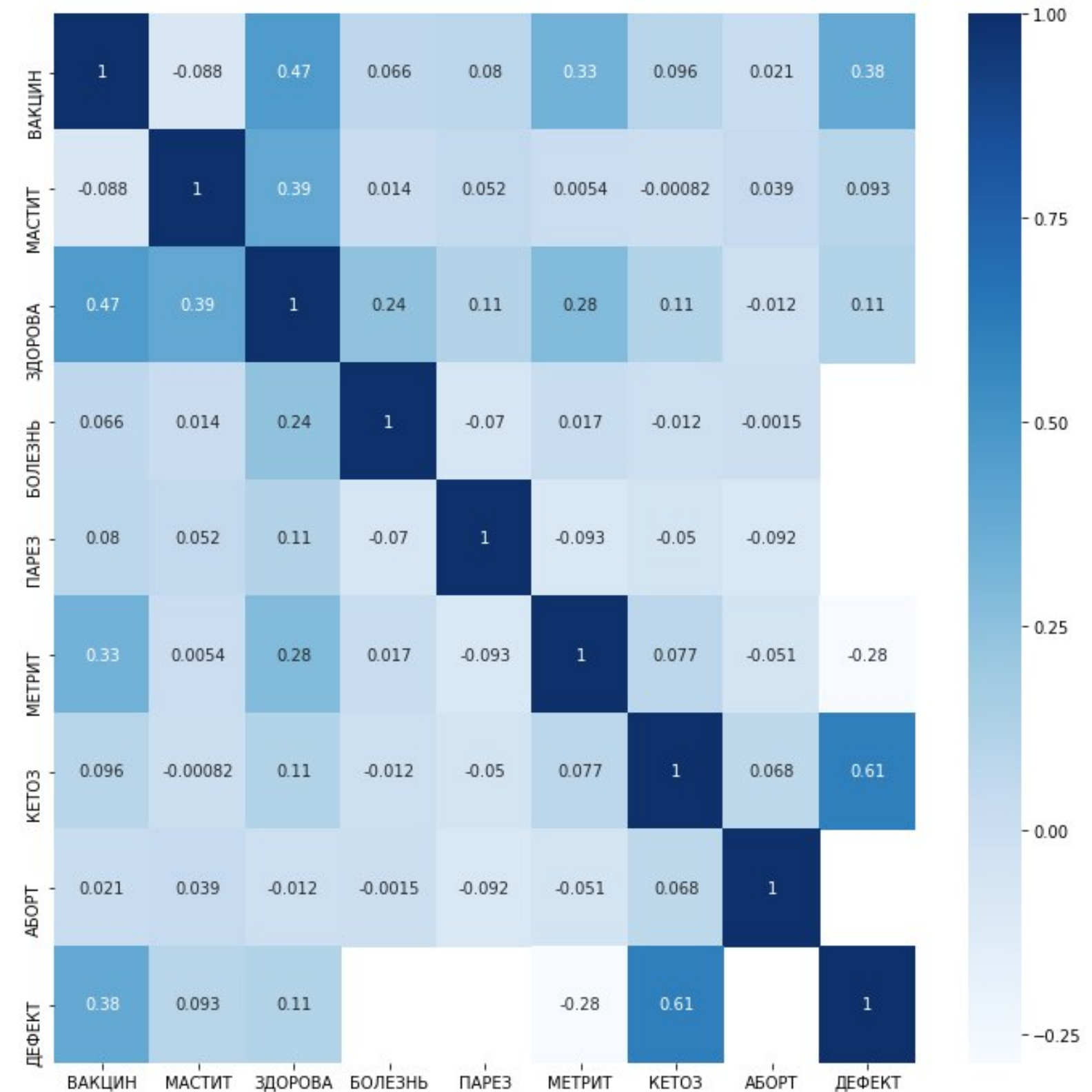
Вывод: В 2019 и 2020 году сезон (зима/лето) влиял на общую заболеваемость - зимой больше, летом - меньше, однако в 2021 и 2022 году объем заболевших особей вырос и сезонность стала практически не влиять

Анализ данных: Взаимосвязи между признаками

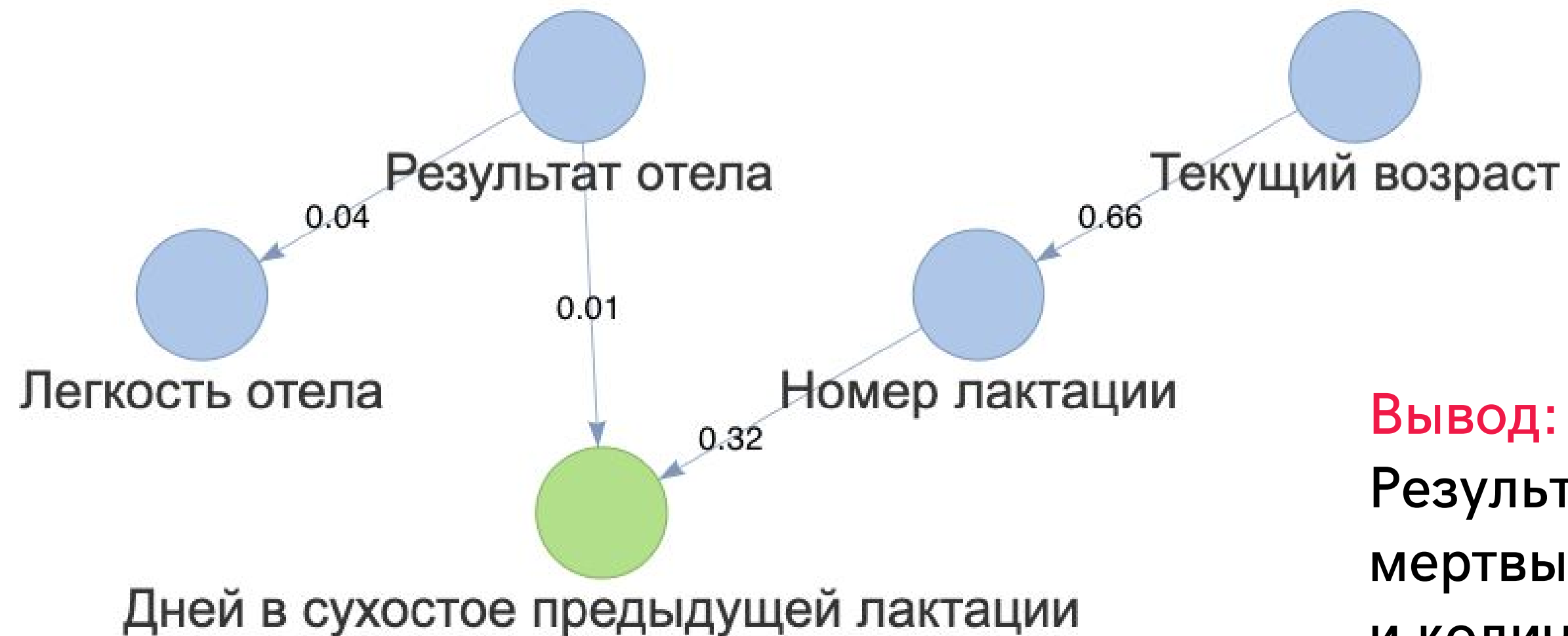
Сделали: Построили матрицу корреляций для выбранных событий писать

Вывод: Между болезнями и другими событиями очень слабая или даже отсутствует линейная корреляция

Идея: Натолкнуло проверить зависимости через другие методы



Матрица корреляции



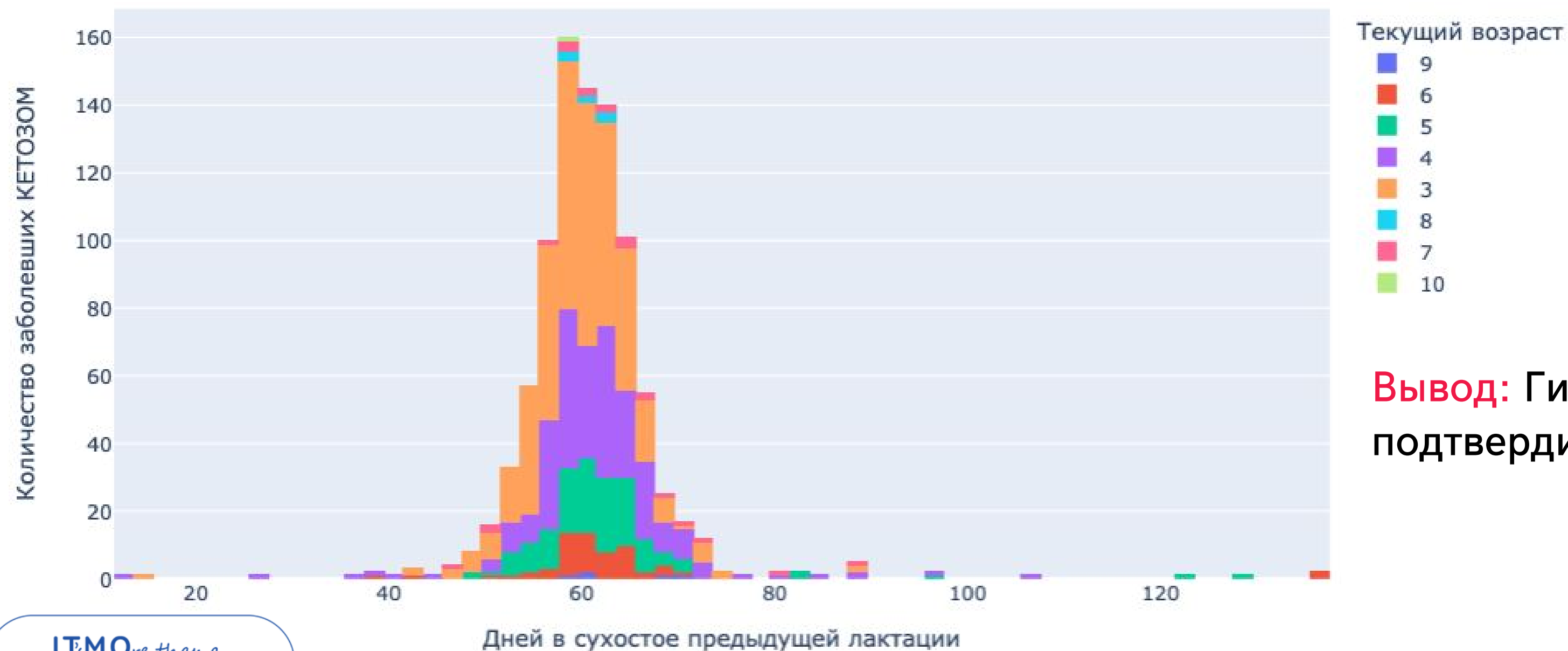
Байесовская сеть связей между событиями,
получен при помощи библиотеки BAMT

Вывод:
Результат отела (живой/
мертвый) влияет на легкость
и количества дней сухостоя,
как и возраст и номер
лактации

Анализ данных: Распределение Кетоза

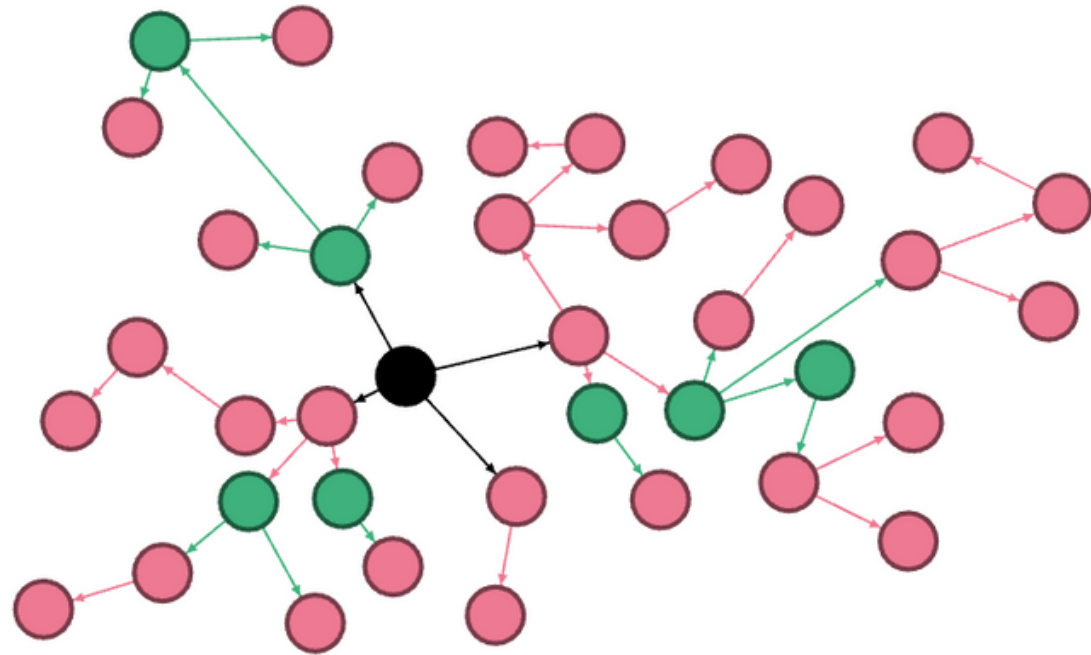
Гипотеза: Заболевание Кетоз зависит от нахождения в сухостое

Распределение заболевших КЕТОЗОМ особей по количеству дней в сухостое

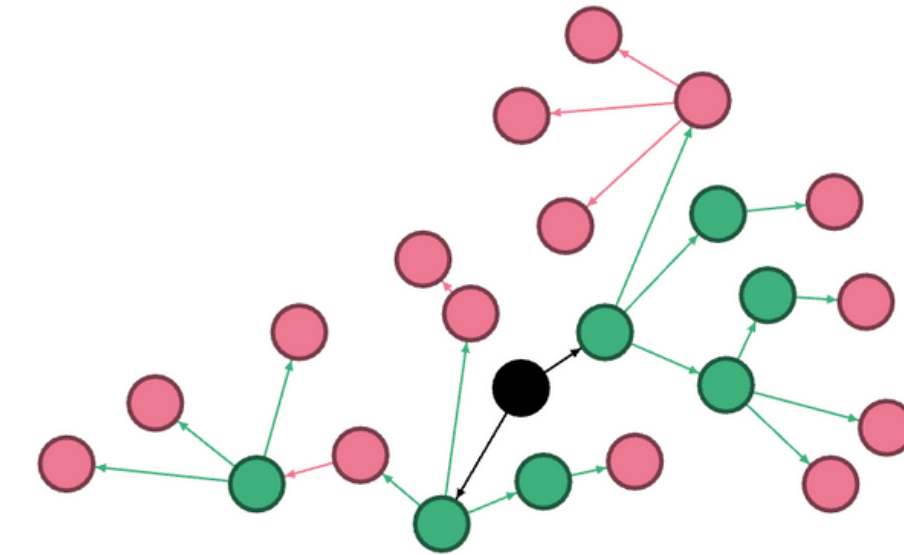


Вывод: Гипотеза не подтвердилась

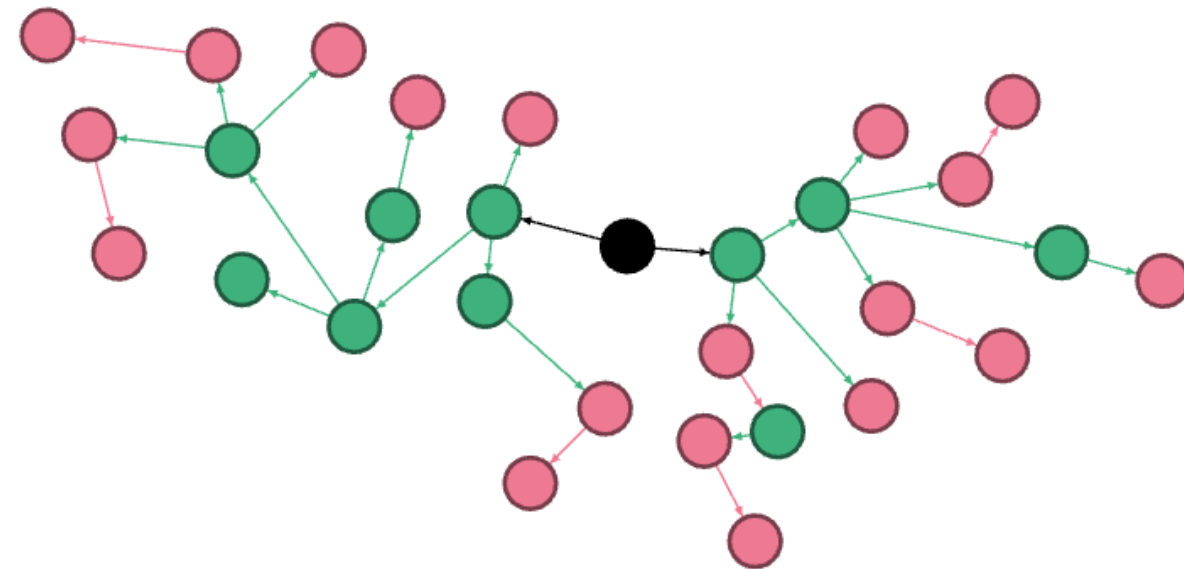
Генеалогический граф



Дерево от матери 3041



Дерево от матери 441



Дерево от матери 483

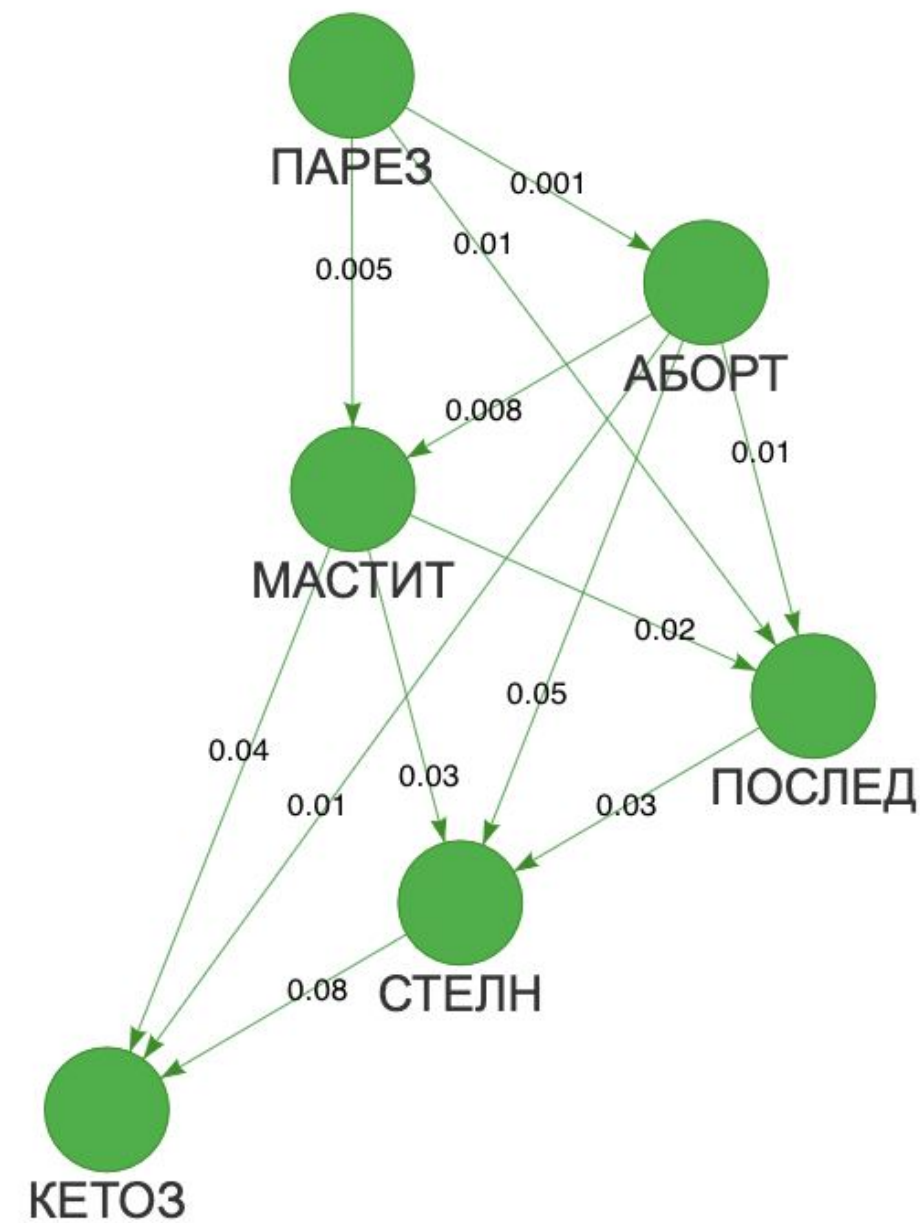
Пояснение:

Зеленый - болела маститом в жизни

Красный - не болела

Визуализация при помощи Gephi

Вывод: Существует связи наследования, но это требует дополнительных исследований



Байесовская сеть связей между событиями, получен при помощи библиотеки BAMT

Пояснение:

Веса для рёбер внутри сети считаются как **условная взаимная информация** (Conditional Mutual Information), что является **взаимной информацией** (Mutual Information) между переменными при условии остальных родителей.

Вывод:

По структуре и связям можно оценить наличие зависимостей между событиями. К примеру, есть зависимость Кетоза от Мастита. Также наблюдается условная зависимость Аборт - Мастит - Стельность. Последнее указывает, что Мастит во время Стельности приводит к Абортам.

	success	fail	prob_success
TM9	237	69	0.77
KM3	896	365	0.71
KM1	1111	668	0.62
KM2	351	212	0.62
KM5	642	454	0.59
KM6	301	207	0.59
KM4	381	334	0.53
KM7	11	24	0.31
KM8	4	16	0.20

Вероятность положительного применения протокола для лечения

Результат: Связь между болезнями и другими событиями очень слабая или даже отсутствует линейная корреляция

Рекомендация: Начинать лечение с KM3, неэффективный KM4

	resistance	age
Номер животного		
512017	{TM9}	7
711136	{KM1}	5

Особь резистентные к протоколам лечения

Результат: Только две живущих ныне коровы обрели резистентности к лекарствам указанных протоколов лечений. Резистентность находили и у других коров в истории.

Вывод: Зависимости между эффективностью протоколов и индивидуальными особенностями коров не существует.



Спасибо за внимание



Bayesian Analytical and
Modelling Toolkit - BAMT
Фреймворк для генерации
Байесовских сетей

FEDOT
Фреймворк для
автоматического
машинного обучения



Стебеньков Андрей
Кейс: Эффективные протоколы лечения коров
Команда: FEDOT HACK TEAM

AgroCode Hack
Пит-стоп #3
18.09.2022