УДК 004.8

DOI 10.25205/1818-7900-2021-19-2-92-101

Предсказание лавинной опасности методами машинного обучения

Н. А. Радеев

*Новосибирский государственный университет
Новосибирск, Россия*

Аннотация

На возникновение снежных лавин главным образом влияют метеорологические условия и конфигурация слоев снежного покрова. Методы машинного обучения имеют предиктивную силу и при должном качестве обучения способны предсказывать новые явления. При обучении моделей машинного обучения на данных о сходах лавин, метеорологических условиях и состоянии снежного покрова получен ансамбль моделей, предсказывающий возможность схода лавины. Представленная в статье модель использует для обучения данные о сходах лавин и метеорологические данные, что позволяет применять полученное решение в большем количестве горных районов, чем решения, использующие более широкий спектр менее доступных данных. Данные о состоянии снежного покрова были сгенерированы программным пакетом SNOWPACK.

Ключевые слова

предсказание снежных лавин, машинное обучение, ансамбль, машина опорных векторов, snowpack, наивный байесовский классификатор, многослойный перцептрон, случайный лес, логистическая регрессия

Для цитирования

Радеев Н. А. Предсказание лавинной опасности методами машинного обучения // Вестник НГУ. Серия: Информационные технологии. 2021. Т. 19, № 2. С. 92–101. DOI 10.25205/1818-7900-2021-19-2-92-101

Avalanches Forecasting Using Machine Learning Methods

N. A. Radeev

*Novosibirsk state university
Novosibirsk, Russian Federation*

Abstract

The occurrence of snow avalanches is mainly influenced by meteorological conditions and the configuration of snow cover layers. Machine learning methods have predictive power and are capable of predicting new events. From the trained machine learning models, an ensemble is obtained that predicts the possibility of avalanches. The model obtained in the article uses avalanche data, meteorological data and generated data on the state of snow cover for training. This allows the resulting solution to be used in more mountainous areas than solutions using a wider range of less available data.

Snow data is generated by the SNOWPACK software package.

Keywords

snow avalanches forecasting, machine learning, ensemble, support vector machine, snowpack, naive bayes, perceptron, random forest, logistic regression

For citation

Radeev N. A. Avalanches Forecasting Using Machine Learning Methods. *Vestnik NSU. Series: Information Technologies*, 2021, vol. 19, no. 2, p. 92–101. (in Russ.) DOI 10.25205/1818-7900-2021-19-2-92-101

© Н. А. Радеев, 2021

ISSN 1818-7900 (Print). ISSN 2410-0420 (Online)

Вестник НГУ. Серия: Информационные технологии. 2021. Том 19, № 2

Vestnik NSU. Series: Information Technologies, 2021, vol. 19, no. 2

Введение

Оценка лавинной опасности является сложной задачей, для решения которой нужно специальное образование и опыт. Также данный процесс требует личного присутствия специалиста в месте, для которого производится оценка. А количество тестов, которые необходимо провести над снежным покровом для получения оценки, приближается к сорока. Всё это делает задачу оценки лавинной опасности очень трудозатратной.

Знание актуального значения оценки лавинной опасности может помочь предпринять превентивные меры по обеспечению защиты благосостояния инфраструктуры, здоровья и жизни людей. В настоящее время системы автоматической оценки лавинной опасности находятся в зачаточном состоянии, и нет решений, пригодных для применения. Таким образом, разработка алгоритма оценки лавинной опасности является актуальной задачей, так как количество специалистов, способных оценивать лавинную опасность, на порядки меньше, чем количество горных районов, в которых необходимо производить указанный вид оценки.

В данной статье описывается разработка хорошо масштабируемого алгоритма предсказания снежных лавин. Последовательно производится обзор существующих решений, описывается, как происходил поиск данных, их очистка, подготовка и предварительный анализ, а также разработка алгоритма оценки лавинной опасности с использованием методов машинного обучения. Обучение и проверка алгоритма происходит на данных из открытых источников, собиравшихся в течение двадцати лет. В их числе метеорологические наблюдения и сгенерированные данные о состоянии снежного покрова.

Обзор существующих работ

Есть множество работ, в которых лавины изучаются с помощью статистических методов, в числе которых и методы машинного обучения. Так, например, в работе [1] строится карта, на которой территории классифицируются как слабо, средне и сильно лавиноопасные на основании статистики о местах схода лавин, их количестве и физических параметров лавин. Было выяснено, что индекс топографической позиции, уклон, осадки и индекс влажности были наиболее полезными данными для обучения алгоритмов. В результате работы получена карта, на основе которой был сделан вывод, что в основном зона лавинной опасности располагается рядом с ручьями и горными склонами возле водных путей.

Все данные собирались из разных источников. В итоге собрано большое разнообразие данных, из которых были получены 14 переменных, используемых в алгоритмах. Данные разбивались в отношении 7 : 3 (7 частей на обучение, 3 на валидацию). Обучение происходило на двух алгоритмах: support vector machine (SVM) и multivariate discriminant analysis (MDA). Выбор объясняется тем, что SVM отлично обобщает и находит паттерны в данных, а MDA предназначен для работы с независимыми переменными, из которых он строит линейные комбинации так, чтобы события из одного класса были на небольшом расстоянии друг от друга.

В работе [2] обучен алгоритм для оценки лавинной опасности по метеорологическим данным, состоянию снежного покрова и другим данным, предоставленным лавинным центром. Модель представляла собой SVM. Данные представляли собой сведения о погоде за три дня и информацию о снежном покрове. Также в результате диалога с лавинным центром был получен список важных индикаторов лавин – около десяти величин, среди которых градиент температуры снега, изменение температуры в предыдущие два дня (любое резкое изменение температуры повышает лавиноопасность) и д. В итоге получилось 44 входных параметра.

Было отобрано 20 наиболее важных параметров. Среди них первые места занимают параметры, которые дал лавинный центр. В свою очередь, некоторые параметры из общедоступного прогноза оказались совсем бесполезными. В общей сложности тренировочные данные

покрывают период в 16 лет – с 1991 по 2007 г. (работа 2008 г., т. е. это максимально свежие данные на тот момент).

В результате работы был получен алгоритм, имеющий метрику точности около 86 % (т. е. среди всех ответов алгоритма 86 % совпали с реальными данными). При этом 28 % прогнозов о возможном сходе лавины (т. е. алгоритм предсказал лавину, но в жизни ее не было) и 8 % прогнозов об отсутствии лавины (алгоритм предсказал отсутствие, а в жизни лавина случилась) являются ложными.

Также существует множество работ, в которых так или иначе проводятся исследования снежного покрова, что имеет косвенное, но очень важное отношение к лавинам. Например, в работе [3] проводится подробный обзор типов снега, апробируется программа SNOWPACK для моделирования состояния снежного покрова, проводится верификация полученных результатов с помощью сравнения с реальными данными, собираемыми в горах Кавказа. А в работе [4] показывается, что использование сгенерированных данных о снежном покрове улучшает предсказания оценки лавинной опасности по пятибалльной шкале.

Развивается направление предсказания лавин с помощью специальных датчиков, устанавливаемых на лавинных склонах [5]. Датчики собирают статистику о колебаниях, возникающих в районе установки. Затем с помощью алгоритмов машинного обучения выделяют те колебания, которые соответствуют сходу лавин, а, например, не пролетевшему мимо самолету.

Обработка данных

Сбор данных

Так как полученное решение должно было эффективно масштабироваться на другие горные районы, то в работе желательно было использование данных, сбор которых не представляет большой сложности и легко автоматизируется. К сожалению, стоит отметить, что сбор статистики по сходу лавин трудоемок. Для этого нужно присутствие специалиста в горном районе, который ежедневно проводит ревизию склонов и фиксирует наличие следов свежих сходов лавин. Тем не менее, стоит ожидать в ближайшем будущем повсеместного появления автоматических детекторов лавин, которые достаточно будет установить в лавинных очагах и периодически собирать с них данные и оказывать техническую поддержку.

Были произведены поиски данных для одного горного района, которые содержали бы статистику схода лавин по дням и еще какие бы то ни было данные о районе, собираемые ежедневно. К счастью, метеорологические станции достаточно широко распространены и выкладывают свои данные в открытый доступ.

В ходе поисков была найдена база данных об окружающей среде Швейцарии. Из этой базы были получены данные о сходах лавин примерно за 20 лет, собираемые практически ежедневно, метеорологические измерения за тот же срок и данные о состоянии снежного покрова, собираемые специалистом примерно раз в две недели. Важным фактором было то, что все три набора данных были собраны в одном горном районе, что гарантировало их совместимость.

Предварительный анализ данных показал, что сведения о снежном покрове собирались недостаточно строго, плохо поддаются манипуляциям ввиду неудачного структурирования измерений и собирались слишком редко. Поэтому они оказались непригодными для использования в решении задачи данной работы.

В результате дальнейших изысканий было решено применить программное генерирование состояния снежного покрова на основе метеорологических измерений, которые собраны в достаточном количестве и были подходящего качества. Для моделирования состояния снежного покрова использован программный пакет SNOWPACK, разработанный швейцарскими исследователями [6]. Данная программа имеет целый ряд подтверждений корректно-

сти своей работы, состоящий из сравнений результатов генерации с данными из реального мира для самых разнообразных мест земного шара.

Очистка и подготовка данных

Так как данные собирались людьми в реальном мире, то в них содержались дефекты вроде пропущенных значений, полей, заполненных не по правилам, и т. п.

Набор сведений о лавинах содержал множество данных: причина схода лавины, тип снега в лавине, размер лавины. В данной работе важен был лишь факт схода либо его отсутствие в течение дня. Поэтому на основе данных был получен список пар (дата, была ли лавина – да / нет).

Метеорологические данные имели пару недостатков. Так как данные начали собирать еще в 1990-е гг., то часть измерений тогда не проводили. В результате не оказалось записей по некоторым показателям за несколько лет. К счастью, эти данные не были нужны для генерации снежного покрова, поэтому от них можно было просто отказаться. Второй проблемой было измерение толщины снега, которое проводилось не слишком часто – раз в две недели. Интуиция подсказывала, что динамика изменения толщины снежного покрова должна играть роль в определении лавинной опасности, поэтому было принято решение разработать алгоритм моделирования данной величины.

Было принято упрощение – за срок в две недели толщина снега либо монотонно убывает, либо монотонно возрастает. В метеорологических данных указывалось количество осадков, выпавших в течение дня. Было решено общий прирост (убыль) толщины H снежного покрова за две недели поделить на общее число осадков P , выпавших в данный срок, – пусть это величина $s = H/P$. Затем недостающие записи восстанавливались пропорционально количеству осадков, выпавших в соответствующий день: если в день k выпало p осадков, то толщина снега в день k увеличилась на $s * p$. Для случая, когда толщина снега за 2 недели убывала, скорость убывания была обратно пропорциональна числу осадков, выпавших в соответствующий день. Таким образом, все недостающие значения были получены.

С данными о состоянии снежного покрова не было никаких проблем, потому что они были получены с помощью генератора. Вследствие этого они сразу были готовы к использованию.

Выделение значимых признаков

Для обучения алгоритма предсказания факта лавин имеющиеся данные нужно было преобразовать к виду (X, y) , где X – вектор фиксированной длины, содержащий данные, описывающие точку в пространстве признаков, а y – целевая бинарная величина, которую алгоритм должен предсказать, получив на вход X .

Метеорологические данные имеют период 30 минут. В свою очередь, статистика по лавинам имеет период 1 день. Другими словами, множество метеорологических данных за сутки нуждалось в агрегации в небольшое количество величин. Например, из температуры были выделены три величины: максимальная температура за сутки, минимальная температура за сутки, разность максимальной и минимальной температуры за сутки. Таким образом, 24 числа были заменены тремя. Подобным образом, исходя из физического смысла показаний и интуитивной оценки важности того или иного показателя для лавинной опасности, были выделены 23 переменные.

Данные о снежном покрове представляют собой численное описание свойств слоев снега: порядковый номер, толщина, высота над уровнем земли, тип снега и другие характеристики. Возникла проблема того, что количество слоев снега – это переменная величина. В начале зимы снега нет совсем – 0 слоев. А в январе количество слоев достигает нескольких сотен. Извлечь полезную информацию из такого массива данных не просто, учитывая, что в целом набор обучающих данных достаточно небольшой. Поэтому из данных о слоях снега были

взяты 11 величин, которые описывают снежный покров в целом, а не каждый слой индивидуально.

Таким образом, было получен вектор X длиной 34, состоящий из вещественных чисел. Все переменные были нормированы до диапазона $[-1; 1]$. Всего в обучающей выборке около 3 000 записей, из них 2 292 об отсутствии лавин и 701 о наличии.

Анализ данных

С помощью статистических функций и библиотек для визуализации данных произведен анализ. Была составлена матрица корреляций признаков, и из пар признаков, имеющих корреляцию больше 0,8, убран один из признаков. Данное действие значительно уменьшает размерность пространства признаков, при этом в данных остается признак, который хранит похожую в смысле корреляции информацию. В результате было удалено 6 признаков, осталось 28.

Далее для каждого признака данные были разбиты по целевой переменной на два множества, и для каждого множества была построена диаграмма размаха (ящик с усами). Таким образом, для каждого признака получились две диаграммы, сообщающие, как распределены данные, где их среднее, максимальное и минимальное значения, выбросы. В результате из обучающей выборки были удалены те признаки, у которых диаграммы размаха для обоих состояний (есть лавина / нет лавины) совпадали. У оставшихся признаков диаграммы заметно разнились. Это говорит о том, что оставшиеся признаки имеют отношение к факту схода лавины. В итоге было удалено 12 признаков, осталось 16.

Разработка алгоритма предсказания

Построение модели

Данные были перемешаны и разбиты на две выборки в отношении 9 : 1. Самая большая выборка нужна для обучения алгоритма (обучающая выборка). Оставшаяся часть данных разбивается на валидационную и тестовую выборки. С помощью валидационной выборки можно подбирать гиперпараметры алгоритма, а на тестовой выборке будут получены итоговые метрики качества.

Были обучены различные алгоритмы классификации: дерево решений, случайный лес деревьев решений, градиентный бустинг, машина опорных векторов, логистическая регрессия, наивный байесовский классификатор, многослойный перцептрон, k ближайших соседей. Каждый алгоритм имеет определенный набор параметров, регулируя которые, можно улучшить или ухудшить метрики качества, полученные на валидационной выборке. Добившись экспериментальным путем наилучших показателей на проверочных и тренировочных данных, алгоритмы были объединены в ансамбль по принципу бэггинга с равными весами. Иначе говоря, каждый алгоритм в ансамбле делал независимое предсказание, а затем итоговое предсказание ансамбля выбиралось простым голосованием по принципу большинства. Так как каждый алгоритм сам по себе имеет некие недостатки, то имеет смысл совместить результаты их работы и выбирать ответ путем голосования среди алгоритмов. Например, один алгоритм хорошо распознавал записи, в которых происходят лавины, но при этом причислял к данному классу и записи, которые на самом деле относятся к классу «нет лавины». А другой алгоритм поступал наоборот. Тогда, усреднив результаты работы обоих алгоритмов, можно было ожидать, что вместе они будут давать более приближенную к правде оценку.

При обучении и проверке моделей машинного обучения использовался ряд метрик: точность, полнота, аккуратность, $f1$ -мера. При этом важно было создать алгоритм, который как можно реже давал бы ложный положительный ответ, – когда в жизни лавина случилась, а алгоритм «сказал», что лавины быть не должно. Таким образом, даже если все остальные метрики были слегка выше, выбирался тот вариант, у которого полнота класса «лавины есть» была больше.

Список испробованных алгоритмов и их лучшие результаты на валидационной выборке приведены в табл. 1.

Таблица 1

Сравнение алгоритмов

Table 1

Comparison of algorithms

№ п/п	Алгоритм	TP	TN	FP	FN	ACC	F1
1	Дерево решений	30	60	26	27	0.63	0.53
2	Градиентный бустинг	31	80	14	25	0.74	0.61
3	Случайный лес деревьев решений	27	89	5	29	0.77	0.61
4	K-ближайших соседей	21	83	11	35	0.69	0.48
5	Машина опорных векторов с полиномиальным ядром 3-го порядка	35	52	59	4	0.58	0.53
6	Машина опорных векторов с линейным ядром	48	54	40	8	0.68	0.67
7	Машина опорных векторов с сигмоидным ядром	59	27	25	29	0.61	0.69
8	Машина опорных векторов с радиально-базисной функцией в качестве ядра	47	54	40	9	0.67	0.66
9	Классификатор на линейном дискриминантном анализе	27	80	14	29	0.71	0.56
10	Наивный байесовский классификатор	42	59	35	14	0.67	0.63
11	Многослойный перцептрон	38	73	21	18	0.74	0.66
12	Логистическая регрессия	50	56	38	6	0.71	0.69

Примечание: TP (True Positive) – лавина была и предсказана; TN (True Negative) – лавины не было, и это предсказано; FP (False Positive) – лавины не было, но она была предсказана; FN (False Negative) – лавина была, но было предсказано, что нет.

В результате комбинирования моделей в ансамбли показатели итогового алгоритма были заметно улучшены. В основе ансамбля легли три версии машины опорных векторов, которые очень редко предсказывали отсутствие лавины для записи, когда на самом деле лавина была. Но в обратную сторону они справлялись не так хорошо и часто предсказывали лавину, когда ее не было. Чтобы компенсировать данный недостаток, в ансамбль были добавлены модели с обратной балансировкой результатов либо модели, в равной степени ошибавшиеся в обоих случаях.

При построении ансамбля основной идеей было взять наиболее «безопасные» модели, т. е. такие, которые меньше всего склонны причислять события из класса «сошла лавина» к классу «лавины не было», и дополнить ансамбль теми моделями, которые могли бы компенсировать недостатки уже взятых в ансамбль и уменьшить ошибку в определении класса «лавины не было».

Был проведен ряд экспериментов по комбинированию моделей, которые хорошо дополняли друг друга (табл. 2). Сравнение производилось на валидационной выборке.

Из полученных ансамблей в первую очередь выбирались те, у которых FN наименьший. Ранее говорилось, что это самая важная метрика, так как предсказать отсутствие лавины тогда, когда она была на самом деле, опасно для жизни людей, и таких случаев следует избегать. Наименьший FN у ансамблей под номерами 2, 5, 6–9. При этом сразу можно отсеять

номера 5–7, так как у них FP больше при тех же значениях FN (рис. 1). Таким образом, остались ансамбли 2, 8, 9. Из них выбор был сделан в пользу ансамбля с наименьшим числом ошибок, т. е. под номером 8. Ансамбль номер 2 хоть и является более «безопасным» в плане предсказаний, метрики качества имеет значительно более плохие. В свою очередь, ансамбль номер 9 дает примерно те же результаты, что и 8. Но так как на тестовой выборке у обоих ансамблей FN одинаковый, а FP меньше у номера 8, то решено остановить выбор на ансамбле номер 8.

Таблица 2

Сравнение ансамблей

Table 2

Comparison of ensembles

№ п/п	Номера моделей, включенных в ансамбль	TP	TN	FP	FN	ACC	F1
1	3, 12	26	89	5	30	0.77	0.6
2	3, 5, 12	50	57	37	6	0.71	0.7
3	3, 5, 11, 12	43	75	19	13	0.79	0.73
4	3, 5, 10, 12	44	70	24	12	0.76	0.71
5	3, 5, 10, 11, 12	49	62	32	7	0.74	0.72
6	3, 5, 6, 10, 11, 12	48	64	30	8	0.75	0.72
7	3, 5, 6, 8, 10, 11, 12	49	58	36	7	0.71	0.7
8	3, 5, 6, 8, 9, 10, 11, 12	48	65	29	8	0.75	0.72
9	3, 5, 6, 7, 8, 9, 10, 11, 12	49	63	31	7	0.75	0.72
10	3, 5	27	89	5	29	0.77	0.61
11	3, 5, 10	46	60	34	10	0.71	0.68
12	2, 3, 5, 10	32	82	24	12	0.76	0.64
13	2, 3, 5, 10, 12	45	64	30	11	0.73	0.69
14	9, 10	23	82	12	33	0.7	0.51
15	5, 9, 10	45	58	36	11	0.69	0.66
16	3, 5, 6, 9, 10	43	67	27	13	0.73	0.68
17	3, 5, 6, 9, 10, 12	43	68	26	13	0.74	0.69

Рис. 1. Диаграмма ложных ответов различных ансамблей
Fig. 1. Diagram of false responses of different ensembles

Нетрудно увидеть (см. табл. 1), что полученный ансамбль имеет меньшее число ошибок, чем все одиночные алгоритмы, кроме случайного леса деревьев решений. Но последний, в свою очередь, является «небезопасным» алгоритмом с очень большим значением FN. Таким образом, полученный ансамбль действительно является наиболее хорошим решением из полученных в данной работе.

В итоговый ансамбль вошли следующие модели:

- 1) Машина опорных векторов с полиномиальным ядром третьего порядка;
- 2) Машина опорных векторов с линейным ядром;
- 3) Машина опорных векторов с радиально-базисной функцией в качестве ядра;
- 4) Логистическая регрессия;
- 5) Классификатор на линейном дискриминантном анализе;
- 6) Случайный лес деревьев решений на 64 деревьях;
- 7) Наивный байесовский классификатор;
- 8) Многослойный перцептрон с двумя скрытыми слоями размерами 64 и 16 нейронов с функцией активации ReLU.

В первых четырех моделях используется балансирование весов классов, так как исходная выборка не сбалансирована (записей из класса «нет лавины» в несколько раз больше записей класса «есть лавина»).

Верификация модели

Проверка работы полученного алгоритма многократно проводилась на валидационной выборке на каждой итерации настройки алгоритма (рис. 2) и один контрольный раз на тестовой выборке. Также были вычислены средние значения метрик (macro avg – среднее арифметическое значение, weighted avg – взвешенное среднее арифметическое пропорционально размерам классов). Разбиение на такие выборки необходимо, чтобы исключить попадание в модель метаинформации, которая неизбежно используется при подгонке параметров модели на валидационной выборке. Поэтому результат, полученный на тестовой выборке (рис. 3), можно считать итоговым, так как модель ни напрямую, ни косвенно не обучалась на данных из тестовой выборки.

а

	precision	recall	f1-score
False	0.89	0.69	0.78
True	0.62	0.86	0.72
accuracy			0.75
macro avg	0.76	0.77	0.75
weighted avg	0.79	0.75	0.76

б

Рис. 2. Матрица предсказаний алгоритма (а) и метрики (б) на валидационной выборке
Fig. 2. Confusion matrix (a) and metrics (b) on validation dataset

Рис. 3. Матрица предсказаний алгоритма (а) и метрики (б) на тестовой выборке
 Fig. 3. Confusion matrix (a) and metrics (b) on test dataset

Заключение

В ходе проделанной работы были собраны, сгенерированы, очищены и приведены к рабочему виду данные. Обучено множество моделей машинного обучения и составлен ансамбль из алгоритмов, имеющий хорошие показатели в предсказании схода лавины по прогнозу погоды. Стоит отметить, что полученное решение использует лишь данные о метеорологической обстановке в горном районе и не использует реальные данные о состоянии снежного покрова и экспертные данные из лавинных центров. Это позволяет использовать полученное решение в горных районах, не имеющих развитых лавинных центров, при условии наличия статистики схода лавин в районе за срок в несколько лет и показаний метеорологических станций. Стоит отметить эффективность генерирования состояния снежного покрова с помощью программы SNOWPACK, которая подтверждается улучшением качества работы алгоритма предсказания лавины при добавлении сгенерированных данных.

В итоге было получено решение, имеющее предиктивную силу. Это свидетельствует о том, что предсказание лавин по метеорологическим данным и сгенерированным данным о состоянии снежного покрова возможно.

Список литературы

1. **Bahram Choubin, Moslem Borji, Amir Mosavi, Farzaneh Sajedi-Hosseini, Vijay P. Singh, Shahaboddin Shamshirband.** Snow avalanche hazard prediction using machine learning methods. *Journal of Hydrology*, 2019, vol. 577, p. 123929. DOI 10.1016/j.jhydrol.2019.123929
2. **Pozdnoukhov A., Purves R. S., Kanevski M.** Applying machine learning methods to avalanche forecasting. *Annals of Glaciology*, 2008, vol. 49. DOI 10.3189/172756408787814870
3. **Марченко Е. С.** Пространственная оценка устойчивости снежного покрова для определения возможности схода лавин разных генетических типов: Автореф. дис. ... канд. геогр. наук / Моск. гос. ун-т. М., 2013.
4. **Schirmer M., Lehning M., Schweizer J.** Statistical forecasting of regional avalanche danger using simulated snow-cover data. *Journal of Glaciology*, 2009, vol. 55, no. 193. DOI 10.3189/002214309790152429
5. **Cordy P., McClung D. M., Hawkins C. J., Tweedy J., Weick T.** Computer assisted avalanche prediction using electronic weather sensor data. DOI 10.1016/j.coldregions.2009.07.006

6. **Радеев Н. А.** Сбор и предобработка данных для системы предсказания лавинной опасности // Тез. докл. Междунар. конф. «Мальцевские чтения 2020». Новосибирск, 2020. С. 91.
7. **Головин Н. А., Савин Н. П., Яхьяева Г. Э.** Применение методов машинного обучения для структурирования базы прецедентов компьютерных атак // Материалы Междунар. конф. «Знания – Онтологии – Теории». Новосибирск, 2019. С. 122–128.

References

1. **Bahram Choubin, Moslem Borji, Amir Mosavi, Farzaneh Sajedi-Hosseini, Vijay P. Singh, Shahaboddin Shamshirband.** Snow avalanche hazard prediction using machine learning methods. *Journal of Hydrology*, 2019, vol. 577, p. 123929. DOI 10.1016/j.jhydrol.2019.123929
2. **Pozdnoukhov A., Purves R. S., Kanevski M.** Applying machine learning methods to avalanche forecasting. *Annals of Glaciology*, 2008, vol. 49. DOI 10.3189/172756408787814870
3. **Marchenko E. S.** Spatial assessment of the stability of snow cover to determine the possibility of avalanches of different genetic types. Author. Dis. ... Cand. Geogr. Sciences. Moscow State University. Moscow, 2013. (in Russ.)
4. **Schirmer M., Lehning M., Schweizer J.** Statistical forecasting of regional avalanche danger using simulated snow-cover data. *Journal of Glaciology*, 2009, vol. 55, no. 193. DOI 10.3189/002214309790152429
5. **Cordy P., McClung D. M., Hawkins C. J., Tweedy J., Weick T.** Computer assisted avalanche prediction using electronic weather sensor data. DOI 10.1016/j.coldregions.2009.07.006
6. **Radeev N. A.** Collection and preprocessing of data for the avalanche hazard prediction system. In: Abstracts of the International Conference “Maltsev Readings 2020”. Novosibirsk, 2020, p. 91. (in Russ.)
7. **Golovin N. A., Savin N. P., Yakhyayeva G. E.** Application of machine learning methods to structure the base of computer attacks precedents. In: Materials of the International Conference “Knowledge – Ontology – Theory”. Novosibirsk, 2019, p. 122–128. (in Russ.)

Материал поступил в редакцию

Received

11.03.2021

Сведения об авторе

Радеев Никита Андреевич, бакалавр кафедры общей информатики факультета информационных технологий Новосибирского государственного университета (Новосибирск, Россия)

n.radeev@g.nsu.ru

Information about the Author

Nikita A. Radeev, Bachelor of General Informatics Department, Faculty of Information Technologies, Novosibirsk state University (Novosibirsk, Russian Federation)

n.radeev@g.nsu.ru