



Proceeding Paper FLOating-Window Projective Separator (FloWPS) Machine Learning Approach to Predict Individual Clinical Efficiency of Cancer Drugs[†]

Nicolas Borisov ¹,*^(D), Victor Tkachev ^{2,3}, Maxim Sorokin ^{2,3}^(D) and Anton Buzdin ^{2,3,4}

- ¹ Moscow Institute of Physics and Technology, 141701 Dolgoprudny, Russia
- ² OmicsWayCorp, Walnut, CA 91788, USA; tkachev@oncobox.com (V.T.); sorokin@oncobox.com (M.S.); buzdin@oncobox.com (A.B.)
- ³ World-Class Research Center "Digital Biodesign and Personalized Healthcare", I.M. Sechenov First Moscow State Medical University, 119991 Moscow, Russia
- ⁴ Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, 117997 Moscow, Russia
- * Correspondence: borisov@oncobox.com; Tel.: +7-903-218-7261
- + Presented at the 1st International Electronic Conference on Biomedicine, Online, 1–26 June 2021; Available online: https://ecb2021.sciforum.net/.

Abstract: (1) Background: Various Machine Learning (ML) methods are applied for the prediction of individual efficiency of cancer treatment regimens. As features for ML, different multi-omics data may be used. We proposed a next-generation ML approach termed FloWPS (FLOating-Window Projective Separator) that filters features before building the ML models to preclude extrapolation in the feature space. (2) Methods: Using Gene Expression Omnibus (GEO), The Cancer Genome Archive (TCGA), and Tumor Alterations Relevant for GEnomics-driven Therapy (TARGET) project databases, as well as our own data, we selected 32 gene expression datasets for cancer patients, annotated with a clinical response status. The biggest dataset included 235 patient cases, and the smallest one had only 41. (3) Results and Discussion: We demonstrated essential improvement of ML quality metrics for FloWPS-based clinical response classifiers for all global ML methods applied, including support vector machines (SVM), random forest (RF), binomial naïve Bayes (BNB), adaptive boosting (ADA), and multi-level perceptron (MLP). Namely, AUC for these classifiers increased from the 0.61–0.88 range to 0.70–0.98. (4) Conclusion: In our ML trial with 32 cancer gene expression datasets, the BNB method with FloWPS showed the best performance, with minimal, median, and maximal AUC values equal to 0.77, 0.86, and 0.98, respectively.

Keywords: bioinformatics; personalized medicine; oncology; chemotherapy; machine learning; omics profiling; support vector machines; random forest; adaptive boosting; multi-level perceptron

1. Background

Machine Learning (ML) methods can offer a wide spectrum of opportunities by nonhypothesis-driven direct linkage of specific molecular features with clinical outcomes, such as responsiveness to certain types of treatment [1–5].

The high-throughput transcriptomic data, including microarray- and next-generation sequencing gene expression profiles, can be utilized for building such classifiers/predictors of clinical response to a certain type of treatment. However, in sharp contrast with the clinical data on COVID-19 [6], acute/chronic liver failure [7], cancer risk [8], and drug repurposing [9], the direct use of ML to personalize the prediction of drug efficiency is problematic. This problem is caused by the lack of sufficient amounts of preceding clinically annotated cases supplemented with high-throughput molecular data (~thousands or tens of thousands of cases per treatment scheme) [10]. As a result, classical ML methods are often not successful in predicting clinical outcomes for several model datasets [11–15].



Citation: Borisov, N.; Tkachev, V.; Sorokin, M.; Buzdin, A. FLOating-Window Projective Separator (FloWPS) Machine Learning Approach to Predict Individual Clinical Efficiency of Cancer Drugs. *Biol. Life Sci. Forum* **2021**, *7*, 23. https://doi.org/10.3390/ ECB2021-10273

Academic Editor: Masaru Tanaka

Published: 31 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). To improve the performance of ML in biomedicine, we recently developed an approach called flexible data trimming (FDT), which removes or excludes extreme values, or outliers from a dataset [2,4,5,16,17]. Excluding non-informative features helps ML classifiers to avoid extrapolation, which is a well-known problem of ML [18–21]. Thus, for every point of a *validation* dataset, the *training* dataset is adjusted to form a floating window. We, therefore, called the respective ML approach *FLOating-Window Projective Separator* (FloWPS) [2,4].

We investigated [4,5] the FloWPS performance for seven popular ML methods, including linear support vector machine (SVM), *k* nearest neighbors (kNN), random forest (RF), Tikhonov (ridge) regression (RR), binomial naïve Bayes (BNB), adaptive boosting (ADA), and multi-layer perceptron (MLP). We performed computational experiments for 32 high-throughput gene expression datasets (41–235 samples per dataset) corresponding to 2596 cancer patients with known responses to chemotherapy treatments [4,5]. We showed that FloWPS essentially improved the classifier quality for all *global* ML methods (SVM, RF, BNB, ADA, MLP), where the AUC for the treatment response classifiers increased from the 0.61–0.88 range to 0.70–0.98 [4,5]. For all the datasets tested, the best performance of FloWPS data trimming was observed for the BNB method, which can be valuable for the further building of ML classifiers in personalized oncology [4,5].

Additionally, to test the robustness of FloWPS-empowered ML methods against overtraining, we interrogated agreement/consensus features between the different ML methods tested, which were used for building mathematical models for the classifiers [4]. The lack of such agreement/consensus could indicate overtraining of the ML classifiers, suggesting random noise instead of extracting significant features distinguishing between the treatment responders and non-responders. If ML methods indeed tend to amplify random noise during overtraining, then one could expect a lack of correlation between the features for geometrically different ML models.

However, we found here that (i) there were statistically significant positive correlations between different ML methods in terms of relative feature importance and (ii) that this correlation was enhanced for the ML methods with FloWPS. We, therefore, conclude that the beneficial role of FloWPS is not due to overtraining [4].

2. Methods

2.1. Clinically Annotated Molecular Datasets

We used 32 publicly available datasets (see Table 1, [3–5]) that contain high-throughput gene expression profiles associated with clinical outcomes of the respective patients [2–5]. Every dataset met the following criteria [3]:

- at least 40 gene expression profiles present;
- data obtained for the same cancer type and using the same experimental platform;
- every profile is linked with the case clinical history;
- all cancers treated with at least one common drug or chemotherapy regimen;
- treatment outcomes are available, enabling the classification of every case as either responder or non-responder.

The dataset preparation for the analysis included the following steps [2,3,5,17]:

- Labeling each patient as either responder or non-responder on the therapy used [3,17];
- For each dataset, finding top marker genes having the highest AUC values for distinguishing responder and non-responder classes [3,17];
- Performing the leave-one-out (LOO) cross-validation procedure to complete the robust core marker gene set used for building the ML model [3,17].

Reference	Dataset ID	Disease Type	Treatment	Experimental Platform	Number NC of Cases (R vs. NR)	Number S of Core Marker Genes
[22,23]	GSE25066	Breast cancer with different hormonal and HER2 status	Neoadjuvant taxane + anthracycline	Affymetrix Human Genome U133 Array	235 (118 R: Complete response + partial response; 117 NR: Residual disease + progressive disease)	20
[24]	GSE41998	Breast cancer with different hormonal and HER2 status	Neoadjuvant doxorubicin + cyclophosphamide, followed by paclitaxel	Affymetrix Human Genome U133 Array	68 (34 R: Complete response + partial response; 34 NR: Residual disease + progressive disease)	11
[25]	GSE9782	Multiple myeloma	Bortezomib monotherapy	Affymetrix Human Genome U133 Array	169 (85 R: Complete response + partial response; 84 NR: No change + progressive disease)	18
[26]	GSE39754	Multiple myeloma	Vincristine + adriamycin + dexamethasone followed by autologous stem cell transplantation (ASCT)	Affymetrix Human Exon 1.0 ST Array	124 (62 R: Complete, near-complete, and very good partial responders, 62 NR: Partial, minor, and worse)	16
[27]	GSE68871	Multiple myeloma	Bortezomib-thalido-mide- dexamethasone	Affymetrix Human Genome U133 Plus	98 (49 R: Complete, near-complete, and very good partial responders, 49 NR: Partial, minor, and worse)	12
[28]	GSE55145	Multiple myeloma	Bortezomib followed by ASCT	Affymetrix Human Exon 1.0 ST Array	56 (28 R: Complete, near-complete, and very good partial responders, 28 NR: Partial, minor. and worse)	14
[5]	https: //www.frontiersin. org/articles/10.3389/ fonc.2021.652063/full# supplementary- material (accessed on 17 November 2021)	Multiple myeloma	Bortezomib, doxorubicin, and dexamethasone (PAD), or bortezomib, cyclophospha-mide, and dexamethasone (VCD)	RNA sequencing, Illumina HiSeq 3000	53 (28 R: complete response + very good partial response; 25 NR: partial response + minimal response)	8
[29,30]	GSE19784_1	Multiple myeloma, ISS stage I	Bortezomib, doxorubicin and dexamethasone (PAD)	Affymetrix Human Genome U133 Plus 2.0 Array	61 (32 R, 29 NR)	7
[29,30]	GSE19784_2	Multiple myeloma, ISS stage II	Bortezomib, doxorubicin and dexamethasone (PAD)	Affymetrix Human Genome U133 Plus 2.0 Array	51 (33 R, 18 NR)	12
[29,30]	GSE19784_3	Multiple myeloma, ISS stage III	Bortezomib, doxorubicin and dexamethasone (PAD)	Affymetrix Human Genome U133 Plus 2.0 Array	41 (29 R, 12 NR)	11
[29,31]	GSE2658	Multiple myeloma	Bortezomib, doxorubicin and dexamethasone (PAD)	Affymetrix Human Genome U133 Plus 2.0 Array	208 (55 R, 153 NR)	16
[32,33]	TARGET-50	Pediatric kidney Wilms tumor	Vincristine sulfate + cyclosporine, cytarabine, daunorubicin + conventional surgery + radiation therapy	Illumina HiSeq 2000	72 (36 R, 36 NR)	14
[32,34]	TARGET-10	Pediatric acute lymphoblastic leukemia	Vincristine sulfate + carboplatin, cyclophosphamide, doxorubicin	Illumina HiSeq 2000	60 (30 R, 30 NR)	14
[32]	TARGET-20	Pediatric acute myeloid leukemia	Non-target drugs (asparaginase, cyclosporine, cytarabine, daunorubicin, etoposide; methotrexate, mitoxantrone), including busulfan and cyclophosphamide	Illumina HiSeq 2000	46 (23 R, 23 NR)	10
[32]	TARGET-20	Pediatric acute myeloid leukemia	Same non-target drugs, but excluding busulfan and cyclophosphamide	Illumina HiSeq 2000	124 (62 R, 62 NR)	16
[35]	GSE18728	Breast cancer	Docetaxel, capecitabine	Affymetrix Human Genome U133 Plus 2.0 Array	61 (23R: Complete response + partial response; 38 NR: Residual disease + progressive disease)	16

Table 1. Clinically annotated gene expression datasets [3–5].

Reference	Dataset ID	Disease Type	Treatment	Experimental Platform	Number NC of Cases (R vs. NR)	Number S of Core Marker Genes
[36,37]	GSE20181	Breast cancer	Letrozole	Affymetrix Human Genome U133A Array	52 (37 R: Complete response + partial response; 15 NR: Residual disease + progressive disease)	11
[38]	GSE20194	Breast cancer	Paclitaxel; (tri)fluoroacetyl chloride; 5-fluorouracil, epirubicin, cyclophosphamide	Affymetrix Human Genome U133A Array	52 (11 R: Complete response + partial response; 41 NR: Residual disease + progressive disease)	10
[39]	GSE23988	Breast cancer	Docetaxel, capecitabine	Affymetrix Human Genome U133A Array	61 (20 R: Complete response + partial response; 41 NR: Residual disease + progressive disease)	18
[40]	GSE32646	Breast cancer	Paclitaxel, 5-fluorouracil, epirubicin, cyclophosphamide	Affymetrix Human Genome U133 Plus 2.0 Array	115 (27 R: Complete response + partial response; 88 NR: Residual disease + progressive disease)	17
[41]	GSE37946	Breast cancer	Trastuzumab	Affymetrix Human Genome U133A Array	50 (27 R: Complete response + partial response; 23 NR: Residual disease + progressive disease)	14
[42]	GSE42822	Breast cancer	Docetaxel, 5-fluorouracil, epirubicin, cyclophosphamide, capecitabine	Affymetrix Human Genome U133A Array	91 (38 R: Complete response + partial response; 53 NR: Residual disease + progressive disease)	13
[43]	GSE5122	Acute myeloid leukemia	Tipifarnib	Affymetrix Human Genome U133A Array	57 (13 R: Complete response + partial response + stable disease; 44 R: Progressive disease)	10
[44]	GSE59515	Breast cancer	Letrozole	Illumina HumanHT-12 V4.0 expression beadchip	75 (51 R: Complete response + partial response; 24 NR: Residual disease + progressive disease)	15
[45]	TCGA-LGG	Low-grade glioma	Temozolomide + (optionally) mibefradil	Illumina HiSeq 2000	131 (100 R: Complete response + partial response + stable disease; 31 NR: Progressive disease)	9
[45]	TCGA-LC	Lung cancer	Paclitaxel + (option- ally),cisplatin/carboplatin, reolysin	Illumina HiSeq 2000	41 (24 R: Complete response + partial response + stable disease; 17 NR: Progressive disease)	7

Table 1. Cont.

2.2. Machine Learning (ML) Application with and without FloWPS

Although modern ML applications in clinical cancer genomics may rely on deep learning methods [46–48], they require large preceding case cohorts [47], which was not the case for either of the gene expression datasets under investigation. Thus, to further characterize them, we used several non-deep ML methods implemented in the Python *sklearn* library [49].

For each ML method, we used a data trimming/preprocessing step using the FloWPS method (R package flowpspkg.tar.gz (Available at https://gitlab.com/borisov_oncobox/flowpspkg, accessed on 17 November 2021)) to increase the robustness and efficiency due to individual sample-specific selection of the training dataset [2,4,17]. A detailed description of the FloWPS algorithm [2,4,17] is given in the Supplementary Material. Among the ML methods, the package *flowpspkg* allows the application of linear/cubic support vector machines (SVM) [2,16,50,51], the *k* nearest neighbors (kNN) method [52], random forest (RF) [53], ridge regression (RR) [54], binomial naïve Bayes (BNB) [55],

adaptive boosting (APA) [11,56,57], and multi-layer perceptron (MLP) [58–60]. For RF, these settings were *n_estimators* = 30, *criterion* = 'entropy'. For BNB: *alpha* = 1.0, *binarize* = 0.0, and *fit_prior* = False. For MLP: *hidden_layer_sizes* = 30, *alpha* = 0.001. To compensate for the possible effect of unequal number of responder and non-responder samples, all SVM and RF calculations were performed by setting *class_weight* = 'balanced' and *class_weight* = 'balanced_subsample'. All other parameters were used with the default settings.

3. Results and Discussion

3.1. Performance of FloWPS with Different Balance Factors for False Positive vs. False Negative Errors

We used FloWPS in combination with seven ML methods, namely, linear support vector machines (SVM) [2,16,50,51], *k* nearest neighbors (kNN) [52], random forest (RF) [53], ridge regression (RR) [54], binomial naïve Bayes (BNB) [55], adaptive boosting (ADA) [11,56,57], and multi-layer perceptron (MLP) [58–60]. For each ML method, we checked the performance of this method with and without FloWPS.

For all ML methods, the FloWPS predictions (P_{Fi}) are equal to the likelihoods for attribution of samples to one of the two classes (clinical responders or non-responders). The *discrimination threshold* (τ), which may be applied to distinguish between the two classes, should be determined according to the cost balance between false positive (FP) and false negative (FN) errors. In our pilot study [2], for the determination of the τ value, we considered the costs for FP and FN errors to be equal and then maximized the overall accuracy rate, ACC = (TP + TN)/(TP + TN + FP + FN), since the class sizes were equal.

In a more general case [4,5], the penalty value $p = B \cdot FP + FN$ is minimized; here, *B* is called the relative balance factor. *B* is less than 1 for the situations when the FN error (e.g., refusal of prescription of a drug that might help the patient) is more dangerous than the FP error (e.g., prescription of a useless treatment). Contrarily, *B* is greater than 1 when it is safer not to prescribe treatment for a patient than to prescribe it. Several practitioners of clinical diagnostic tests have different opinions on how high/low this balancing factor should be. In different applications, the preferred values can be B = 4 [61–63], B < 0.16 [64], 4.5 < B < 5 [44], B < 5 [65], B > 10 for emergency medicine only [66], B > 5 for toxicology [67].

In the case of oncological disease, *B* should be low when only *one or few* treatment options is/are available for a certain patient because the refusal to give a treatment may cause serious harm to the patient. Contrarily, in the situation when the doctor must select the best treatment plan among *multiple* options available, the risk of a wrong drug prescription will be higher, and *B* should be high as well. For our analyses, we used five model settings of *B* equal to 0.1, 0.25, 1, 4, and 10.

For the quality metrics of AUC, sensitivity (Sn), and specificity (Sp), see Table 2 and Figures 1 and 2. We found that the use of FloWPS has considerably improved the AUC metric for all global ML methods investigated (SVM, RF, BNB, ADA, and MLP) but had no effect on the performance of local methods, kNN and RR. For the global ML methods, FloWPS improved the classifier quality and increased AUC from the 0.65–0.85 range to 0.78–0.98, and AUC median values—from the 0.70–0.77 range to 0.76–0.82 (Table 2).

Table 2. Performance metrics for seven ML methods with default settings for datasets with equal numbers of responders and non-responders [4].

ML Method	Method Type	Median AUC without FloWPS	Median AUC with FloWPS	Paired <i>t</i> -Test <i>p</i> -Value for AUC with-vsw/o FloWPS	Advantage of FloWPS	Median Sn at <i>B</i> = 4	Median Sp at <i>B</i> = 0.25
SVM	Global	0.74	0.80	$1.3 imes 10^{-5}$	Yes	0.45	0.42
kNN	Local	0.76	0.75	0.53	No	0.25	0.34
RF	Global	0.74	0.82	$1.3 imes10^{-5}$	Yes	0.45	0.42
RR	Local	0.80	0.79	0.16	No	0.36	0.41
BNB	Global	0.77	0.82	$2.7 imes10^{-4}$	Yes	0.51	0.58
ADA	Global	0.70	0.76	$2.4 imes10^{-4}$	Yes	0.32	0.41
MLP	Global	0.73	0.82	$6.4 imes10^{-5}$	Yes	0.53	0.53

Yes-FloWPS is beneficial for ML quality, No-FloWPS is not beneficial for ML quality.



Figure 1. Receiver-operator curves (ROC) showing the dependence of sensitivity (*Sn*) upon specificity (*Sp*) for FloWPS-based classifier of treatment response for datasets with *core marker genes* [2]. Red dots: confidence parameter p = 0.95, blue dots: p = 0.90. Panels represent different clinically annotated datasets, (**A**): GSE25066 [22,23]; (**B**): GSE41998 [24]; (**C**): GSE9782 [25]; (**D**): GSE39754 [26]; (**E**): GSE68871 [27]; (**F**): GSE55134 [28]; (**G**): TARGET-50 [32,33]; (**H**): TARGET-10 [32,34]; (**I**,J): TARGET-20 [32] with busulfan and cyclophosphamide, respectively.



Figure 2. The area under the receiver-operator curve (AUC), sensitivity (Sn), and Specificity (Sp) for five ML methods (BNB, MLP, RF, RR, and SVM), both with (red line) and without (blue line) FloWPS, during the classification of response to PAD/VCD treatment of 53 MM patients (full cohort). Parameter B is the false positive vs. false negative balance factor [5].

Table 2 and Figure 2 summarize these findings. Considering the quality criterion that combines the highest AUC, the highest Sn at B = 4, and the highest Sp at B = 0.25, the top three methods were BNB, MLP, and RF. Figure 2 shows how the AUC, Sn, and Sp values depend on different values of B. For all ML methods, the application of FloWPS increased the quality of the classifiers built, as reflected by the AUC metric (Table 2, Figure 2). Taking the three criteria of AUC, Sn, and Sp together, the optimal solution was provided by the BNB method with FloWPS (AUC = 0.84) for the multiple myeloma full cohort (Figure 2).

3.2. Correlation Study between Different ML Methods at the Level of Feature Importance

To check whether the FloWPS enhancement of the different ML is caused by overtraining of mathematical models or whether FloWPS helps to extract the essential characteristics in the feature space that separate responders from non-responders, we performed the analysis of the feature importance.

For linear SVM, RF, RR, BNB, and MLP methods, and for all transcriptomic datasets tested, we calculated the *relative importance*, I_f , of each gene expression feature f in the dataset, using the following attributes of ML classes in Python library *sklearn* [61]:

For linear SVM: $I_f = |coef_[0]_f|$, where $coef_[0]$ is the normal vector to the separation hyperplane between responders and non-responders in the feature space in the training model.

For RF, $I_f = |feature_importances_f|$ from the training model.

For RR, $I_f = \sum_t |X_fit_{tf}|$, where the summation runs through every sample *t* in the training model.

For BNB, $I_f = \sum_c feature_count_{cf}$, where the values named *feature_count_{cf}* denote the number of samples encountered for each class *c* and feature *f* during the fitting of the training model.

For MLP, $I_f = \sum_t |coefs[0]_{tf}|$, where $coefs[0]_{tf}$ is the coefficient matrix in the first layer of the neural network for feature *f* of sample *t* in the training model.

For each validation point *I*, the I_f was averaged over all predication-accountable set S_i . We showed positive pairwise correlations between the different ML methods at the level of the relative importance of different features tested (Table 3, Figure 3).

Greater similarities between I_f marks for the different ML methods reflect more robust applications of the ML. Importantly, the correlations for the ML methods with FloWPS were always higher than for the methods without FloWPS. This suggests the beneficial role of FloWPS in extracting informative features from noisy data. In this model, the biggest similarity was observed for the pair of RR and BNB methods.

Table 3. Median pairwise (with/without FloWPS) correlations between deferent ML methods at feature (gene expression) importance (I_f) level. Figures above main diagonal: Pearson correlations; figures below: Spearman correlations [4].

	SVM	RF	RR	BNB	MLP
SVM	1	0.53/0.34	0.40/0.19	0.37/0.24	0.46/0.33
RF	0.55/0.40	1	0.51/0.35	0.48/0.33	0.590.40
RR	0.39/0.14	0.32/0.04	1	0.93/0.88	0.89/0.76
BNB	0.34/0.14	0.31/0.09	0.79/0.64	1	0.81/0.61
MLP	0.46/0.30	0.38/0.17	0.52/0.06	0.46/0.12	1

3.3. Discussion

The application of ML methods for the prescription of certain drugs to certain patients in omics-based bioinformatics is limited by the deficiency of annotated case histories, where we know the responses to the treatment combined with the omics (gene expression, mutation, methylation, phosphorylation, etc.) profiles. These annotated cases should be collected for a certain combination of the cancer type/localization and treatment regimen.

However, the accuracy for the ML-based classification may be low upon the blind cross-validation of the ML modes since many ML methods, designed for global separation of different classes of points in the feature space, are prone to overtraining when the number of preceding cases is low. Global ML methods may also fail if there is only local rather than global order in the placement of different classes in the feature space [2,18–20].



Figure 3. Pairwise correlations (red—with FloWPS, green—without FloWPS) at feature (gene expression) level between different ML methods: figures above the main diagonal—Pearson correlation coefficient, figures below—Spearman [4].

To improve the performance of the global ML methods, we suggested a novel approach some years ago [2,4,17]. It includes some elements of the local methods, e.g., using the flexible data trimming that avoids extrapolation in the feature space for each validation point, followed by selecting only several nearest neighbors from the training dataset. In such a way, the whole ML classifier becomes a hybrid, both global and local [2,4,17].

According to this hybrid approach, for each validation point, the training of ML models is performed in the individually tailored feature space. Every validation point is surrounded by *a floating window* from the points of the training dataset, and the irrelevant features are avoided using the rectangular projections in the feature space. That is why the approach was called FLOating-Window Projective Separator, FloWPS [2,4,17].

Overtraining, together with extrapolation, is also a well-known Achilles heel of ML. We, therefore, tested if FloWPS helps to extract truly significant features or if it simply adapts to random noise, thus, causing overfitting. We compared four global ML methods (SVM, RF, BNB, and MLP) and one local ML method (RR) to check similarities between them in terms of the relative importance of distinct individual features. We confirmed that all these five ML methods were positively correlated at the level of feature importance (Table 3, Figure 3). Moreover, using FloWPS significantly enhanced such correlations in all the cases examined (Table 3, Figure 3). These results clearly suggest that FloWPS is helpful for extracting relevant information, and it does not merely follow the random noise [4].

Overall, we suggest that using correlations between different ML methods at the level of the relative feature importance may be used as an evaluation metric of ML suitability for building classifiers utilizing omics data. In this case, the higher the correlation is, the



bigger the probability that the separation of responders from non-responders is robust and non-overtrained should be (Figure 4).

Figure 4. Correlation values between different ML methods at the feature (gene expression) level for each dataset (1—GSE25066 [22,23]; 2—GSE41998 [24]; 3—GSE9782 [25]; 4—GSE39754 [26]; 5—GSE68871 [27]; 6—GSE55134 [28]; 7—TARGET-50 [32,33]; 8—TARGET-10 [32,34]; 9—TARGET-20 [32] with busulfan; 10—TARGET-20 [32] with cyclophosphamide; 11—GSE18728 [35]; 12—GSE20181 [36,37]; 13—GSE20194 [38]; 14—GSE23988 [39]; 15—GSE32646 [40]; 16—GSE37946 [41]; 17—GSE42822 [42]; 18—GSE5122 [43]; 19—GSE59515 [44]; 20—TCGA-LGG [45]; 21—TCGA-LC [45]) are shown through the horizontal axis). (A)—Pearson, FloWPS, (B)—Pearson, no FloWPS, (C)—Spearman, FloWPS, (D)—Spearman, no FloWPS [4].

Many gene expression datasets are too small to be used as training datasets in personalized oncology. However, there is a possibility to aggregate different smaller datasets into bigger ones. For such aggregation, one may use our new Shambhala-1 [68] and Shahmhala-2 [69] harmonizing techniques, which are capable of merging arbitrary numbers of datasets obtained using arbitrary experimental platforms.

To our knowledge, Shambhala-1/2 were the first uniformly shaped gene expression harmonizers to ever be applied to process the RNAseq and microarray data together [68,69]. We hope that Shambhala-1/2 will become useful for the broad spectrum of applications for combining different expression datasets. One of its major strengths is the stability of the final output data, which is not biased/distorted upon each new round of harmonization.

However, one should keep in mind that both Shambhala-1/2 and FloWPS approaches are algorithmically complex and time-consuming/resource-demanding. Therefore, parallel execution of the program code using both these approaches may be advantageous [4,69].

4. Conclusions

Machine Learning (ML) in personalized medicine can be characterized as an a posteriori paradigm since it relies on the association of molecular and clinical features in the training dataset, which may be used for artificial intelligence (AI) applications [2–4]. Yet, the number of preceding cases in omics-based personalized oncology, e.g., gene expression profiles equipped with known responses to certain types of cancer treatment, is still generally insufficient for the application of conventional ML methods to predict individual clinical responses.

The deficiency of preceding cases for personalized drug prescription in oncology leads to extrapolation in the feature space when building the mathematical models for ML, and the extrapolation leads to model overtraining, which decreases the prediction accuracies. To overcome the problem above, we suggested a novel approach termed FLOating-Window Projective Separator, FloWPS [2,4,5].

By using the FloWPS paradigm, the whole ML classifier becomes a hybrid, both global and local. Namely, for each validation point, training of the ML models is performed in the individually tailored feature space. Moreover, every validation point is surrounded by a floating window from the points of the training dataset, and the irrelevant features are avoided using the rectangular projections in the feature space [2,4,5].

In our ML trial with 32 cancer gene expression datasets, the BNB method with FloWPS showed the best performance, with minimal, median, and maximal AUC values equal to 0.77, 0.86, and 0.98, respectively [4].

Our FloWPS method is unprecedented because it does not use any pre-selected analytical form of transformation kernels but instead adapts the feature space heuristically for every particular validation case. The success of using FloWPS for the real-world gene expression datasets, including tens to hundreds of samples, prompts further trials of its applicability in biomedicine and in other fields where increased accuracy of ML classifiers is needed [2,4,5].

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ECB2021-10273/s1.

Author Contributions: Testing and debugging the computational code, V.T.; identification of relevant gene expression datasets, M.S.; design of the research and preparation of the paper, A.B. and N.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financed by the Ministry of Science and Higher Education of the Russian Federation within the framework of state support for the creation and development of World-Class Research Centers "Digital biodesign and personalized healthcare #075-15-2022-304".

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: See Table 1.

Acknowledgments: The cloud-based computational facilities were sponsored by Amazon and Microsoft Azure grants.

Conflicts of Interest: Authors declare no conflict of interest. The funder had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- Buzdin, A.; Sorokin, M.; Poddubskaya, E.; Borisov, N. High-Throughput Mutation Data Now Complement Transcriptomic Profiling: Advances in Molecular Pathway Activation Analysis Approach in Cancer Biology. *Cancer Inform.* 2019, 18, 117693511983884. [CrossRef] [PubMed]
- Tkachev, V.; Sorokin, M.; Mescheryakov, A.; Simonov, A.; Garazha, A.; Buzdin, A.; Muchnik, I.; Borisov, N. Floating-Window Projective Separator (FloWPS): A Data Trimming Tool for Support Vector Machines (SVM) to Improve Robustness of the Classifier. *Front. Genet.* 2019, 9, 717. [CrossRef] [PubMed]

- 3. Borisov, N.; Sorokin, M.; Tkachev, V.; Garazha, A.; Buzdin, A. Cancer gene expression profiles associated with clinical outcomes to chemotherapy treatments. *BMC Med. Genom.* **2020**, *13*, 111. [CrossRef] [PubMed]
- Tkachev, V.; Sorokin, M.; Borisov, C.; Garazha, A.; Buzdin, A.; Borisov, N. Flexible Data Trimming Improves Performance of Global Machine Learning Methods in Omics-Based Personalized Oncology. *Int. J. Mol. Sci.* 2020, 21, 713. [CrossRef]
- Borisov, N.; Sergeeva, A.; Suntsova, M.; Raevskiy, M.; Gaifullin, N.; Mendeleeva, L.; Gudkov, A.; Nareiko, M.; Garazha, A.; Tkachev, V.; et al. Machine Learning Applicability for Classification of PAD/VCD Chemotherapy Response Using 53 Multiple Myeloma RNA Sequencing Profiles. *Front. Oncol.* 2021, 11, 652063. [CrossRef]
- 6. Chadaga, K.; Prabhu, S.; Umakanth, S.; Bhat, K.; Sampathila, N.; Chadaga, R. COVID-19 Mortality Prediction among Patients using Epidemiological parameters: An Ensemble Machine Learning Approach. *Eng. Sci.* **2021**, *16*, 221–233. [CrossRef]
- Musunuri, B.; Shetty, S.; Shetty, D.K.; Vanahalli, M.K.; Pradhan, A.; Naik, N.; Paul, R. Acute-on-Chronic Liver Failure Mortality Prediction using an Artificial Neural Network. *Eng. Sci.* 2021, 15, 187–196. [CrossRef]
- 8. Khalsan, M.; Machado, L.R.; Al-Shamery, E.S.; Ajit, S.; Anthony, K.; Mu, M.; Agyeman, M.O. A Survey of Machine Learning Approaches Applied to Gene Expression Analysis for Cancer Prediction. *IEEE Access* **2022**, *10*, 27522–27534. [CrossRef]
- 9. Cong, Y.; Shintani, M.; Imanari, F.; Osada, N.; Endo, T. A New Approach to Drug Repurposing with Two-Stage Prediction, Machine Learning, and Unsupervised Clustering of Gene Expression. *OMICS J. Integr. Biol.* **2022**, *26*, 339–347. [CrossRef]
- 10. Azarkhalili, B.; Saberi, A.; Chitsaz, H.; Sharifi-Zarchi, A. DeePathology: Deep Multi-Task Learning for Inferring Molecular Pathology from Cancer Transcriptome. *Sci. Rep.* **2019**, *9*, 16526. [CrossRef]
- 11. Turki, T.; Wang, J.T.L. Clinical intelligence: New machine learning techniques for predicting clinical drug response. *Comput. Biol. Med.* **2019**, *107*, 302–322. [CrossRef] [PubMed]
- 12. Turki, T.; Wei, Z. A link prediction approach to cancer drug sensitivity prediction. *BMC Syst. Biol.* 2017, *11*, 94. [CrossRef] [PubMed]
- Turki, T.; Wei, Z. Learning approaches to improve prediction of drug sensitivity in breast cancer patients. In Proceedings of the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016; pp. 3314–3320.
- 14. Turki, T.; Wei, Z.; Wang, J.T.L. Transfer Learning Approaches to Improve Drug Sensitivity Prediction in Multiple Myeloma Patients. *IEEE Access* 2017, *5*, 7381–7393. [CrossRef]
- 15. Turki, T.; Wei, Z.; Wang, J.T.L. A transfer learning approach via procrustes analysis and mean shift for cancer drug sensitivity prediction. *J. Bioinform. Comput. Biol.* **2018**, *16*, 1840014. [CrossRef] [PubMed]
- Borisov, N.; Tkachev, V.; Suntsova, M.; Kovalchuk, O.; Zhavoronkov, A.; Muchnik, I.; Buzdin, A. A method of gene expression data transfer from cell lines to cancer patients for machine-learning prediction of drug efficiency. *Cell Cycle* 2018, 17, 486–491. [CrossRef]
- 17. Borisov, N.; Buzdin, A. New Paradigm of Machine Learning (ML) in Personalized Oncology: Data Trimming for Squeezing More Biomarkers From Clinical Datasets. *Front. Oncol.* **2019**, *9*, 658. [CrossRef]
- 18. Arimoto, R.; Prasad, M.-A.; Gifford, E.M. Development of CYP3A4 inhibition models: Comparisons of machine-learning techniques and molecular descriptors. *J. Biomol. Screen.* **2005**, *10*, 197–205. [CrossRef]
- 19. Balabin, R.M.; Lomakina, E.I. Support vector machine regression (LS-SVM): An alternative to artificial neural networks (ANNs) for the analysis of quantum chemistry data? *Phys. Chem. Chem. Phys.* **2011**, *13*, 11710–11718. [CrossRef]
- 20. Balabin, R.M.; Smirnov, S.V. Interpolation and extrapolation problems of multivariate regression in analytical chemistry: Benchmarking the robustness on near-infrared (NIR) spectroscopy data. *Analyst* **2012**, *137*, 1604–1610. [CrossRef]
- 21. Betrie, G.D.; Tesfamariam, S.; Morin, K.A.; Sadiq, R. Predicting copper concentrations in acid mine drainage: A comparative analysis of five machine learning techniques. *Environ. Monit. Assess.* **2013**, *185*, 4171–4182. [CrossRef]
- Hatzis, C.; Pusztai, L.; Valero, V.; Booser, D.J.; Esserman, L.; Lluch, A.; Vidaurre, T.; Holmes, F.; Souchon, E.; Wang, H.; et al. A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. *JAMA* 2011, 305, 1873–1881. [CrossRef] [PubMed]
- Itoh, M.; Iwamoto, T.; Matsuoka, J.; Nogami, T.; Motoki, T.; Shien, T.; Taira, N.; Niikura, N.; Hayashi, N.; Ohtani, S.; et al. Estrogen receptor (ER) mRNA expression and molecular subtype distribution in ER-negative/progesterone receptor-positive breast cancers. *Breast Cancer Res. Treat.* 2014, 143, 403–409. [CrossRef] [PubMed]
- Horak, C.E.; Pusztai, L.; Xing, G.; Trifan, O.C.; Saura, C.; Tseng, L.-M.; Chan, S.; Welcher, R.; Liu, D. Biomarker analysis of neoadjuvant doxorubicin/cyclophosphamide followed by ixabepilone or Paclitaxel in early-stage breast cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 2013, 19, 1587–1595. [CrossRef] [PubMed]
- Mulligan, G.; Mitsiades, C.; Bryant, B.; Zhan, F.; Chng, W.J.; Roels, S.; Koenig, E.; Fergus, A.; Huang, Y.; Richardson, P.; et al. Gene expression profiling and correlation with outcome in clinical trials of the proteasome inhibitor bortezomib. *Blood* 2007, 109, 3177–3188. [CrossRef] [PubMed]
- Chauhan, D.; Tian, Z.; Nicholson, B.; Kumar, K.G.S.; Zhou, B.; Carrasco, R.; McDermott, J.L.; Leach, C.A.; Fulcinniti, M.; Kodrasov, M.P.; et al. A small molecule inhibitor of ubiquitin-specific protease-7 induces apoptosis in multiple myeloma cells and overcomes bortezomib resistance. *Cancer Cell* 2012, 22, 345–358. [CrossRef] [PubMed]
- 27. Terragna, C.; Remondini, D.; Martello, M.; Zamagni, E.; Pantani, L.; Patriarca, F.; Pezzi, A.; Levi, G.; Offidani, M.; Proserpio, I.; et al. The genetic and genomic background of multiple myeloma patients achieving complete response after induction therapy with bortezomib, thalidomide and dexamethasone (VTD). *Oncotarget* **2016**, *7*, 9666–9679. [CrossRef]

- Amin, S.B.; Yip, W.K.; Minvielle, S.; Broyl, A.; Li, Y.; Hanlon, B.; Swanson, D.; Shah, P.K.; Moreau, P.; Van Der Holt, B.; et al. Gene expression profile alone is inadequate in predicting complete response in multiple myeloma. *Leukemia* 2014, 28, 2229–2234. [CrossRef]
- 29. Ubels, J.; Sonneveld, P.; van Beers, E.H.; Broijl, A.; van Vliet, M.H.; de Ridder, J. Predicting treatment benefit in multiple myeloma through simulation of alternative treatment effects. *Nat. Commun.* **2018**, *9*, 2943. [CrossRef]
- Broyl, A.; Hose, D.; Lokhorst, H.; de Knegt, Y.; Peeters, J.; Jauch, A.; Bertsch, U.; Buijs, A.; Stevens-Kroef, M.; Beverloo, H.B.; et al. Gene expression profiling for molecular classification of multiple myeloma in newly diagnosed patients. *Blood* 2010, 116, 2543–2553. [CrossRef]
- 31. Zhan, F.; Huang, Y.; Colla, S.; Stewart, J.P.; Hanamura, I.; Gupta, S.; Epstein, J.; Yaccoby, S.; Sawyer, J.; Burington, B.; et al. The molecular classification of multiple myeloma. *Blood* **2006**, *108*, 2020–2028. [CrossRef]
- 32. Goldman, M.; Craft, B.; Swatloski, T.; Cline, M.; Morozova, O.; Diekhans, M.; Haussler, D.; Zhu, J. The UCSC Cancer Genomics Browser: Update 2015. *Nucleic Acids Res.* 2015, *43*, D812–D817. [CrossRef] [PubMed]
- Walz, A.L.; Ooms, A.; Gadd, S.; Gerhard, D.S.; Smith, M.A.; Guidry Auvil, J.M.; Meerzaman, D.; Chen, Q.-R.; Hsu, C.H.; Yan, C.; et al. Recurrent DGCR8, DROSHA, and SIX Homeodomain Mutations in Favorable Histology Wilms Tumors. *Cancer Cell* 2015, 27, 286–297. [CrossRef] [PubMed]
- Tricoli, J.V.; Blair, D.G.; Anders, C.K.; Bleyer, W.A.; Boardman, L.A.; Khan, J.; Kummar, S.; Hayes-Lattin, B.; Hunger, S.P.; Merchant, M.; et al. Biologic and clinical characteristics of adolescent and young adult cancers: Acute lymphoblastic leukemia, colorectal cancer, breast cancer, melanoma, and sarcoma: Biology of AYA Cancers. *Cancer* 2016, 122, 1017–1028. [CrossRef] [PubMed]
- Korde, L.A.; Lusa, L.; McShane, L.; Lebowitz, P.F.; Lukes, L.; Camphausen, K.; Parker, J.S.; Swain, S.M.; Hunter, K.; Zujewski, J.A. Gene expression pathway analysis to predict response to neoadjuvant docetaxel and capecitabine for breast cancer. *Breast Cancer Res. Treat.* 2010, 119, 685–699. [CrossRef] [PubMed]
- Miller, W.R.; Larionov, A. Changes in expression of oestrogen regulated and proliferation genes with neoadjuvant treatment highlight heterogeneity of clinical resistance to the aromatase inhibitor, letrozole. *Breast Cancer Res. BCR* 2010, 12, R52. [CrossRef] [PubMed]
- 37. Miller, W.R.; Larionov, A.; Anderson, T.J.; Evans, D.B.; Dixon, J.M. Sequential changes in gene expression profiles in breast cancers during treatment with the aromatase inhibitor, letrozole. *Pharm. J.* **2012**, *12*, 10–21. [CrossRef]
- Popovici, V.; Chen, W.; Gallas, B.G.; Hatzis, C.; Shi, W.; Samuelson, F.W.; Nikolsky, Y.; Tsyganova, M.; Ishkin, A.; Nikolskaya, T.; et al. Effect of training-sample size and classification difficulty on the accuracy of genomic predictors. *Breast Cancer Res. BCR* 2010, 12, R5. [CrossRef]
- Iwamoto, T.; Bianchini, G.; Booser, D.; Qi, Y.; Coutant, C.; Shiang, C.Y.-H.; Santarpia, L.; Matsuoka, J.; Hortobagyi, G.N.; Symmans, W.F.; et al. Gene pathways associated with prognosis and chemotherapy sensitivity in molecular subtypes of breast cancer. *J. Natl. Cancer Inst.* 2011, 103, 264–272. [CrossRef]
- Miyake, T.; Nakayama, T.; Naoi, Y.; Yamamoto, N.; Otani, Y.; Kim, S.J.; Shimazu, K.; Shimomura, A.; Maruyama, N.; Tamaki, Y.; et al. GSTP1 expression predicts poor pathological complete response to neoadjuvant chemotherapy in ER-negative breast cancer. *Cancer Sci.* 2012, 103, 913–920. [CrossRef]
- Liu, J.C.; Voisin, V.; Bader, G.D.; Deng, T.; Pusztai, L.; Symmans, W.F.; Esteva, F.J.; Egan, S.E.; Zacksenhaus, E. Seventeen-gene signature from enriched Her2/Neu mammary tumor-initiating cells predicts clinical outcome for human HER2+:ERα- breast cancer. *Proc. Natl. Acad. Sci. USA* 2012, 109, 5832–5837. [CrossRef]
- Shen, K.; Qi, Y.; Song, N.; Tian, C.; Rice, S.D.; Gabrin, M.J.; Brower, S.L.; Symmans, W.F.; O'Shaughnessy, J.A.; Holmes, F.A.; et al. Cell line derived multi-gene predictor of pathologic response to neoadjuvant chemotherapy in breast cancer: A validation study on US Oncology 02-103 clinical trial. *BMC Med. Genom.* 2012, *5*, 51. [CrossRef] [PubMed]
- Raponi, M.; Harousseau, J.-L.; Lancet, J.E.; Löwenberg, B.; Stone, R.; Zhang, Y.; Rackoff, W.; Wang, Y.; Atkins, D. Identification of molecular predictors of response in a study of tipifarnib treatment in relapsed and refractory acute myelogenous leukemia. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 2007, 13, 2254–2260. [CrossRef] [PubMed]
- Turnbull, A.K.; Arthur, L.M.; Renshaw, L.; Larionov, A.A.; Kay, C.; Dunbier, A.K.; Thomas, J.S.; Dowsett, M.; Sims, A.H.; Dixon, J.M. Accurate Prediction and Validation of Response to Endocrine Therapy in Breast Cancer. J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol. 2015, 33, 2270–2278. [CrossRef] [PubMed]
- 45. Tomczak, K.; Czerwi?ska, P.; Wiznerowicz, M. The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. *Contemp. Oncol. Pozn. Pol.* **2015**, *19*, A68–A77. [CrossRef]
- 46. Yuan, Y.; Shi, Y.; Li, C.; Kim, J.; Cai, W.; Han, Z.; Feng, D.D. DeepGene: An advanced cancer type classifier based on deep learning and somatic point mutations. *BMC Bioinform.* **2016**, *17*, 476. [CrossRef]
- 47. Yuan, Y.; Shi, Y.; Su, X.; Zou, X.; Luo, Q.; Feng, D.D.; Cai, W.; Han, Z.-G. Cancer type prediction based on copy number aberration and chromatin 3D structure with convolutional neural networks. *BMC Genom.* **2018**, *19*, 565. [CrossRef]
- Huang, Z.; Johnson, T.S.; Han, Z.; Helm, B.; Cao, S.; Zhang, C.; Salama, P.; Rizkalla, M.; Yu, C.Y.; Cheng, J.; et al. Deep learning-based cancer survival prognosis from RNA-seq data: Approaches and evaluations. *BMC Med. Genom.* 2020, 13, 41. [CrossRef]
- 49. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

- Bartlett, P.; Shawe-taylor, J. Generalization Performance of Support Vector Machines and Other Pattern Classifiers. Adv. Kernel Methods Support Vector Learn. 1999, 43–54.
- 51. Vapnik, V.; Chapelle, O. Bounds on Error Expectation for Support Vector Machines. Neural Comput. 2000, 12, 2013–2036. [CrossRef]
- 52. Altman, N.S. An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. Am. Stat. 1992, 46, 175–185. [CrossRef]
- 53. Toloşi, L.; Lengauer, T. Classification with correlated features: Unreliability of feature ranking and solutions. *Bioinformatics* **2011**, 27, 1986–1994. [CrossRef] [PubMed]
- 54. Tikhonov, A.; Yakovlevich Arsenin, V. Solutions of Ill-Posed Problems; Springer: Berlin/Heidelberg, Germany, 1977.
- 55. Webb, G.I.; Boughton, J.R.; Wang, Z. Not So Naive Bayes: Aggregating One-Dependence Estimators. *Mach. Learn.* 2005, 58, 5–24. [CrossRef]
- 56. Wang, Z.; Yang, H.; Wu, Z.; Wang, T.; Li, W.; Tang, Y.; Liu, G. In Silico Prediction of Blood-Brain Barrier Permeability of Compounds by Machine Learning and Resampling Methods. *ChemMedChem* **2018**, *13*, 2189–2201. [CrossRef] [PubMed]
- Yosipof, A.; Guedes, R.C.; García-Sosa, A.T. Data Mining and Machine Learning Models for Predicting Drug Likeness and Their Disease or Organ Category. Front. Chem. 2018, 6, 162. [CrossRef] [PubMed]
- Prados, J.; Kalousis, A.; Sanchez, J.-C.; Allard, L.; Carrette, O.; Hilario, M. Mining mass spectra for diagnosis and biomarker discovery of cerebral accidents. *Proteomics* 2004, *4*, 2320–2332. [CrossRef]
- 59. Marvin, M.; Seymour, A. Papert Perceptrons—Expanded Edition: An Introduction to Computational Geometry; MIT Press: Boston, MA, USA, 1987.
- 60. Robin, X.; Turck, N.; Hainard, A.; Lisacek, F.; Sanchez, J.-C.; Müller, M. Bioinformatics for protein biomarker panel classification: What is needed to bring biomarker panels into in vitro diagnostics? *Expert Rev. Proteom.* **2009**, *6*, 675–689. [CrossRef]
- 61. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Müller, A.; Nothman, J.; Louppe, G.; et al. Scikit-learn: Machine Learning in Python. *arXiv* 2012, arXiv:12010490 Cs.
- 62. Ioannidis, J.P.A.; Hozo, I.; Djulbegovic, B. Optimal type I and type II error pairs when the available sample size is fixed. *J. Clin. Epidemiol.* **2013**, *66*, 903–910. [CrossRef]
- 63. Wetterslev, J.; Jakobsen, J.C.; Gluud, C. Trial Sequential Analysis in systematic reviews with meta-analysis. *BMC Med. Res. Methodol.* **2017**, *17*, 39. [CrossRef]
- 64. Kim, H.-Y. Statistical notes for clinical researchers: Type I and type II errors in statistical decision. *Restor. Dent. Endod.* **2015**, *40*, 249. [CrossRef] [PubMed]
- 65. Lu, J.; Qiu, Y.; Deng, A. A note on Type S/M errors in hypothesis testing. Br. J. Math. Stat. Psychol. 2019, 72, 1–17. [CrossRef]
- 66. Cummins, R.O.; Hazinski, M.F. Guidelines based on fear of type II (false-negative) errors : Why we dropped the pulse check for lay rescuers. *Circulation* **2000**, *102*, I377–I379. [CrossRef] [PubMed]
- 67. Rodriguez, P.; Maestre, Z.; Martinez-Madrid, M.; Reynoldson, T.B. Evaluating the Type II error rate in a sediment toxicity classification using the Reference Condition Approach. *Aquat. Toxicol.* **2011**, *101*, 207–213. [CrossRef] [PubMed]
- Borisov, N.; Shabalina, I.; Tkachev, V.; Sorokin, M.; Garazha, A.; Pulin, A.; Eremin, I.I.; Buzdin, A. Shambhala: A platform-agnostic data harmonizer for gene expression data. *BMC Bioinform.* 2019, 20, 66. [CrossRef]
- 69. Borisov, N.; Sorokin, M.; Zolotovskaya, M.; Borisov, C.; Buzdin, A. Shambhala-2: A Protocol for Uniformly Shaped Harmonization of Gene Expression Profiles of Various Formats. *Curr. Protoc.* **2022**, *2*, e444. [CrossRef]